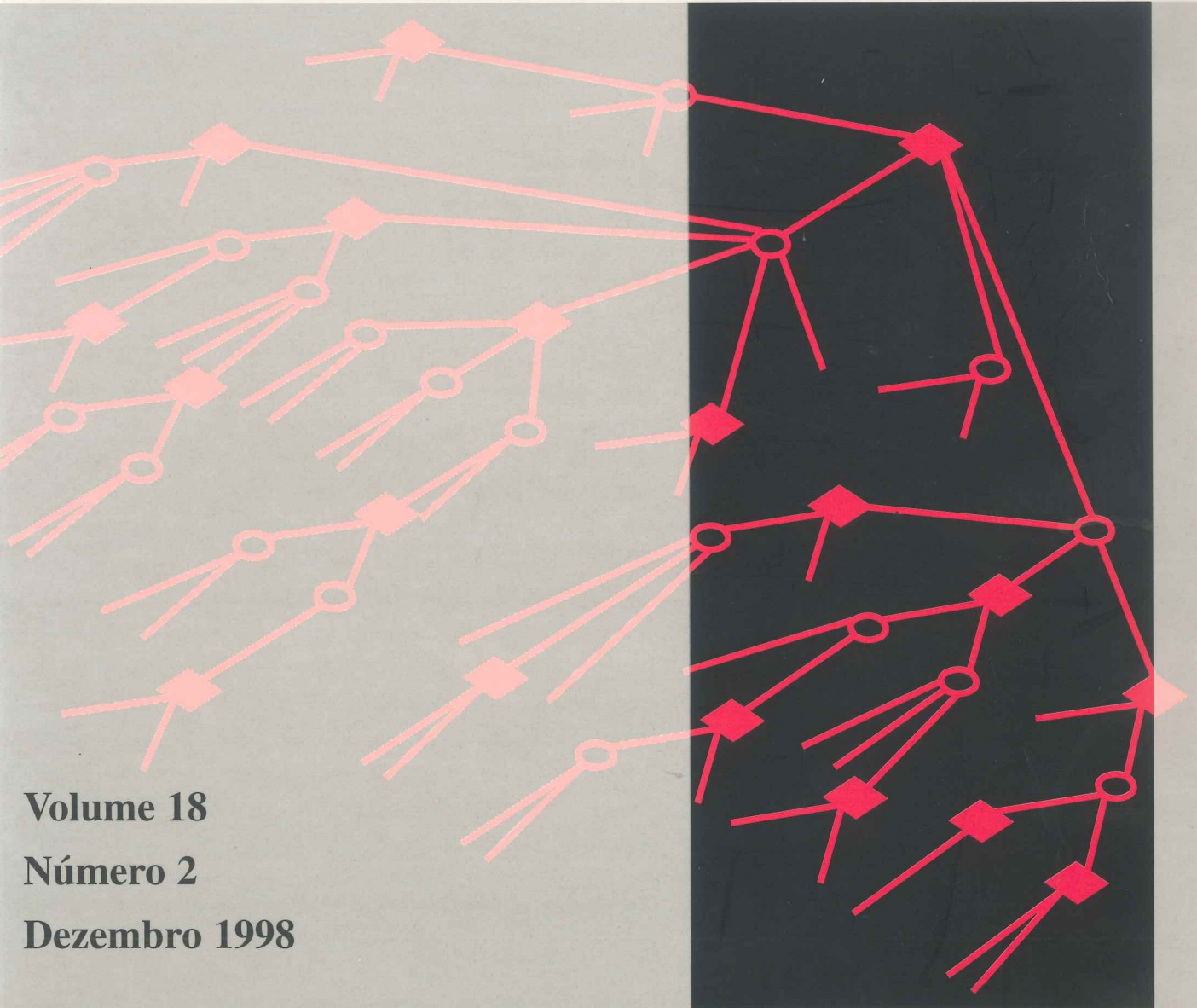


Investigação Operacional



Volume 18

Número 2

Dezembro 1998

INVESTIGAÇÃO OPERACIONAL

Propriedade:

APDIO — Associação Portuguesa de Investigação Operacional

ESTATUTO EDITORIAL

<<Investigação Operacional>>, órgão oficial da APDIO cobre uma larga gama de assuntos reflectindo assim a grande diversidade de profissões e interesses dos sócios da Associação, bem como as muitas áreas de aplicação da I. O. O seu objectivo primordial é promover a aplicação do método e técnicas da I.O. aos problemas da Sociedade Portuguesa.

A publicação acolhe contribuições nos campos da metodologia, técnicas, e áreas de aplicação e software de I. O. sendo no entanto dada prioridade a bons casos de estudo de carácter eminentemente prático.

Distribuição gratuita aos sócios da APDIO

INVESTIGAÇÃO OPERACIONAL

Volume 18 - nº 2 - Dezembro 1998

Publicação Semestral

Editor Principal: Joaquim J. Júdice
Universidade de Coimbra

Comissão Editorial

M. Teresa Almeida
Inst. Sup. Economia e Gestão

Laureano Escudero
IBM, Espanha

J. Pinto Paixão
Univ. de Lisboa

Jaime Barceló
Univ. de Barcelona

J. Soeiro Ferreira
Univ. do Porto

M. Vaz Pato
Inst. Sup. Economia e Gestão

Paulo Barcia
Univ. Nova de Lisboa

J. Fernando Gonçalves
Univ. do Porto

Celso Ribeiro
Univ. Católica, Rio Janeiro

Isabel Branco
Univ. de Lisboa

Clóvis Gonzaga
Univ. Fed., Rio Janeiro

A. Guimarães Rodrigues
Univ. do Minho

António Câmara
Univ. Nova de Lisboa

Luís Gouveia
Univ. de Lisboa

Mário S. Rosa
Univ. de Coimbra

C. Bana e Costa
Inst. Superior Técnico

Rui C. Guimarães
Univ. do Porto

J. Pinho de Sousa
Univ. do Porto

M. Eugénia Captivo
Univ. de Lisboa

J. Assis Lopes
Inst. Superior Técnico

Reinaldo Sousa
Univ. Católica, Rio Janeiro

Jorge O. Cerdeira
Inst. Sup. de Agronomia

N. Maculan
Univ. Fed., Rio Janeiro

L. Valavares Tavares
Inst. Superior Técnico

João Clímaco
Univ. de Coimbra

Ernesto Q. Martins
Univ. de Coimbra

Isabel H. Themido
Inst. Superior Técnico

J. Dias Coelho
Univ. Nova de Lisboa

Vladimiro Miranda
Univ. do Porto

B. Calafate Vasconcelos
Univ. do Porto

J. Rodrigues Dias
Univ. de Évora

José M. Viegas
Inst. Superior Técnico

A Revista “INVESTIGAÇÃO OPERACIONAL” está registada na Secretaria de Estado da Comunicação Social sob o nº 108335.

Esta Revista é distribuída gratuitamente aos sócios da APDIO. As informações sobre inscrições na Associação, assim como a correspondência para a Revista devem ser enviadas para a sede da APDIO - Associação Portuguesa de Investigação Operacional - CESUR, Instituto Superior Técnico, Av. Rovisco Pais, 1049 - 001 Lisboa.

Este Volume foi subsidiado por:

Fundação para a Ciência e Tecnologia

Fundação Calouste Gulbenkian

Para efeitos de dactilografia e composição, foram utilizados equipamentos gentilmente postos à disposição pelo Centro de Investigação Operacional (DEIO - FCUL).

Assinatura: 5.000\$00

UMA HEURÍSTICA NEURONAL PARA A ELABORAÇÃO DE HORÁRIOS ESCOLARES

Paulo Carrasco

Escola Superior de Gestão, Hotelaria e Turismo
Universidade do Algarve

Margarida Vaz Pato

Instituto Superior de Economia e Gestão
Universidade Técnica de Lisboa
Rua do Quelhas, 6
1200 Lisboa - Portugal

Abstract

The paper analyses the problem of timetabling in high schools. This slow and laborious task is often performed manually, through a process of trial and error. A heuristic based on Hopfield's Artificial Neural Network is proposed for this problem. Furthermore, the computational results obtained from one illustrative timetabling instance and four pseudo-real timetabling problems are discussed. This paper ends with some conclusions about this heuristic approach.

Resumo

A elaboração de horários escolares é uma tarefa complexa e morosa, desenvolvida periodicamente na maioria das instituições de ensino. Neste artigo é proposta uma heurística baseada numa rede neuronal de Hopfield para Problemas de Elaboração de Horários Escolares característicos de escolas de ensino secundário. Referem-se os resultados computacionais obtidos com uma pequena instância exemplificativa e quatro cenários pseudo-reais, terminando-se o artigo com algumas conclusões.

Keywords

Timetabling Problem, Hopfield Neural Network, Heuristics.

Introdução

O processo de construção de horários escolares é desenvolvido em todas as instituições de ensino pelo menos uma vez em cada ano. Levada a cabo por técnicos especializados, esta tarefa morosa decorre num contexto de "tentativa e erro" com base na experiência adquirida e no conhecimento das características da instituição.

A modelização deste problema no campo da Investigação Operacional tem originado diferentes abordagens (Carter (1995), de Werra (1985)), correspondendo normalmente a problemas de optimização complexos, com elevada dimensão e de difícil resolução prática.

* Este trabalho foi parcialmente suportado pelo Centro de Investigação Operacional (FCUL) ao abrigo do PRAXIS XXI (JNICT) Projecto nr. 2/2.1/MAT/139/94.

O renovado interesse pela teoria das redes neuronais artificiais, em particular da rede neuronal de Hopfield, evidenciou algumas potencialidades destas na resolução de problemas de optimização complexos. No entanto, poucas aplicações desta metodologia na elaboração de horários têm surgido na literatura. Em Yu (1990) é apresentada uma heurística neuronal para um problema de horários muito simplificado. Por sua vez, Gislén, Peterson e Söderberg (1992) utilizaram com sucesso um modelo neuronal de Potts para a criação de horários escolares.

Neste artigo, estuda-se uma heurística baseada numa rede neuronal de Hopfield para o Problema de Elaboração de Horários Escolares (P.E.H.E.). Na secção 2 introduzem-se sumariamente os elementos básicos de uma rede neuronal particularizando para o modelo de Hopfield. Na secção 3 estuda-se o P.E.H.E., descrevendo-se as suas principais características no contexto das instituições de ensino secundário em Portugal e, a partir de uma formalização em Programação Não Linear Binária (P.N.L.B.), apresenta-se a respectiva formulação em termos neuronais. A descrição do algoritmo e a definição dos parâmetros são desenvolvidos na secção 4. Na secção 5 mostram-se alguns resultados obtidos com uma pequena instância exemplificativa e com alguns P.E.H.E. pseudo-reais. A secção 6 termina o artigo com algumas conclusões e considerações relevantes.

2. Redes de Neurónios Artificiais. O Modelo Neuronal de Hopfield

O interesse pelo estudo da teoria das redes neuronais e sua modelização em termos computacionais é devido a McCulloch e Pitts (1943) que pela primeira vez apresentaram um modelo de um neurónio artificial. Desde então vários foram os desenvolvimentos desta teoria publicados em numerosos trabalhos das mais diversas áreas (Inteligência Artificial, Ciências da Computação e Investigação Operacional). Tal facto deve-se a um conjunto de características particulares normalmente associadas às redes neuronais artificiais, nomeadamente o processamento de forma distribuída, a capacidade de aprendizagem e a possibilidade de adaptação a contextos de incerteza e elevada complexidade.

A abordagem de problemas de optimização através de redes neuronais foi introduzida por Hopfield (1982, 1984) e posteriormente por Hopfield e Tank (1985) mostrando que através de uma modelização adequada é possível transpor problemas de optimização discreta para uma estrutura neuronal, na qual se torna possível o cálculo de soluções de qualidade não necessariamente óptimas.

A rede neuronal de Hopfield é formalmente constituída por processadores (neurónios) com dois estados possíveis na versão discreta (1 - activo, 0 - inactivo). Neste artigo é utilizada a versão contínua do modelo de Hopfield dado o seu melhor ajustamento ao problema em questão.

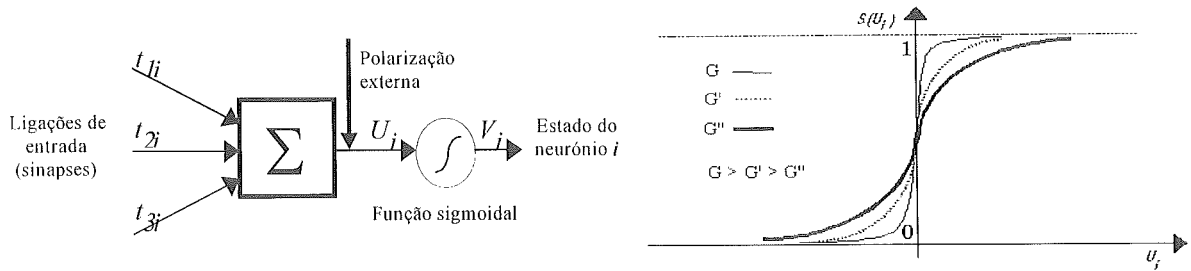


Figura 1 - Esquema de um neurônio artificial de Hopfield

Às ligações entre neurónios (sinapses) são associadas ponderações expressas na matriz simétrica $[t_{ki}]_{k,i = 1, \dots, N}$, em que t_{ki} representa a ponderação da ligação existente entre o neurónio k e o neurónio i . Se $t_{ki} > 0$ a sinapse diz-se excitadora e se $t_{ki} < 0$ a sinapse considera-se inibidora.

Cada neurónio encontra-se densamente conectado aos restantes, funcionando individualmente do seguinte modo (ver Figura 1):

- inicialmente o neurónio i efectua uma soma ponderada dos estados dos restantes neurónios que lhe estão conectados;
- de seguida é determinado o valor U_i através da adição de uma polarização externa, I_i , à soma anterior;
- por último, o estado do neurónio i , V_i , é determinado através da aplicação de uma função sigmoideal, $S(U_i)$. Esta função reproduz o padrão de resposta de um neurónio biológico face a um estímulo, para valores do sinal U_i negativos V_i aproxima-se de 0 (estado inactivo), enquanto que para U_i positivos V_i tende para 1 (estado activo).

Assim, na entrada do neurónio i apresenta-se um sinal dado pela expressão:

$$U_i = \sum_{k=1}^N t_{ki} V_k + I_i \tag{1}$$

verificando-se à saída do neurónio i um estado dado pelo valor da função sigmoideal na qual G representa um parâmetro, factor de ganho (Figura 1):

$$V_i = S(U_i) = \frac{1}{2} (1 + \tanh(GU_i)) \tag{2}$$

Considerando uma rede com N neurónios, a sua evolução temporal é representada em cada momento τ por um vector caracterizador dos estados dos neurónios $V(\tau) = (V_1(\tau), V_2(\tau), \dots, V_N(\tau))$. Em τ , o potencial total energético concentrado na rede neuronal é dado pela função de energia:

$$E(V(\tau)) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N t_{ij} V_i(\tau) V_j(\tau) - \sum_{i=1}^N I_i V_i(\tau) + \sum_{i=1}^N \frac{1}{R_i} \int_0^{V_i} S_i^{-1}(V) dV \tag{3}$$

sendo R_i um parâmetro representativo da resistência de entrada no neurónio i . Esta função tem um papel fundamental na resolução de problemas de optimização através de redes neuronais.

Hopfield (1984) provou que a função de energia é uma função de Lyapunov, atendendo a que a matriz de ponderações associada às ligações sinápticas é simétrica. Provou ainda que, utilizando um processo particular de actualização dos estados dos neurónios da rede, a energia decresce ao longo do tempo. Consequentemente, a função de energia tende a estabilizar num ponto de estacionaridade, o qual pode corresponder a um mínimo energético. A evolução dinâmica da rede é expressa pelas seguintes equações de actualização dos estados dos neurónios:

$$\frac{dU_i(\tau)}{d\tau} = -\frac{U_i(\tau)}{K} + \sum_{k=1}^N t_{ki}V_k(\tau) + I_i \quad (4)$$

na qual, K , a constante de perda energética à entrada do neurónio, toma os valores propostos por Wilson e Pawley (1988). Estas equações diferenciais não lineares traduzem a capacidade minimizadora da rede de Hopfield e são resolvidas na prática pelo método aproximado de Euler (Braun (1975)).

As etapas para a modelização do problema de horários escolares em termos de uma rede neuronal de Hopfield, e os aspectos essenciais do seu funcionamento são apresentados na próxima secção.

3. Problema de Elaboração de Horários Escolares e o Modelo Neuronal

O problema de elaboração de horários escolares presente nas instituições de ensino portuguesas, nomeadamente nas secundárias, possui características que diferem substancialmente do relativo às escolas dos EUA e Canadá (Carrasco (1995)). De uma forma geral, o processo de construção de horários pode ser dividido em quatro fases distintas.

- 1ª Fase - **Inscrição de cada aluno numa das áreas de ensino e escolha das optativas** - No nosso país a escolha de uma área de ensino condiciona o aluno à frequência de um curriculum de cadeiras fundamentalmente fixo com poucas optativas, ao contrário do que sucede no sistema americano onde o curriculum é maioritariamente opcional.
- 2ª Fase - **Elaboração de turmas formadas pelos alunos com curriculum comum** - Para atender às cadeiras opcionais as turmas podem ser divididas ou fundidas para formar turmas conjuntas de alunos com optativas iguais. A dimensão das turmas está também sujeita a um limite máximo.
- 3ª Fase - **Afectação das turmas aos professores** - De acordo com a área de especialização e o número de horas de aulas semanais programadas para cada professor é-lhe atribuído um conjunto de cadeiras e turmas para leccionar. Esta correspondência professor/cadeira/turma será designada por disciplina. Assim, se o professor 1 leccionar Matemática à turma A, tal afectação será representada pela disciplina MATEMÁTICA-1A. Nas optativas, por exemplo, se o professor 2 leccionar Inglês aos alunos provenientes das turmas A e B, a disciplina será INGLÊS-2AB.

4ª Fase - Elaboração dos horários - Esta última fase consiste propriamente na construção de horários e é sobre ela que nos vamos concentrar ao longo do artigo.

O P.E.H.E. insere-se na 4ª Fase e consiste na criação de horários semanais para professores e alunos. Neste contexto, a resolução do problema passa pela elaboração de um conjunto de afectações entre grupos definidos de disciplinas (constituídas na 3ª Fase), salas e períodos de tempo, ou seja, os horários. Os condicionalismos inerentes a uma situação real deste tipo são normalmente numerosos e complexos, podendo ser agrupados de acordo com o seu grau de obrigatoriedade.

Condicionalismos de 1º nível - Estes asseguram a satisfação das seguintes restrições:

- uma aula de uma disciplina ocupa exactamente uma sala, e um só período de tempo;
- cada professor só lecciona no máximo uma disciplina, em cada tempo;
- cada aluno assiste no máximo a uma disciplina, em cada tempo;
- o número de tempos por semana para cada cadeira é cumprido, de acordo com o plano da área de ensino do aluno;
- a capacidade das salas e a sua adequação às disciplinas é verificada;
- devem ser respeitadas algumas considerações pedagógicas, tais como, a obrigatoriedade de existência de um intervalo de tempo para almoço e de um limite máximo diário de aulas por disciplina.

Condicionalismos de 2º nível - Neste grupo pode ser incluído um vasto grupo de exigências relativas a aspectos desejáveis na solução, tanto para os professores como para os alunos. Consideraremos apenas os seguintes aspectos:

- a ocorrência de furos deverá ser evitada, bem como a dispersão das aulas ao longo de toda a semana;
- as preferências horárias manifestadas pelos professores devem ser respeitadas.

Esta caracterização do problema de horários cobre apenas as particularidades mais relevantes da elaboração de horários na realidade de uma escola secundária portuguesa, podendo ser adicionadas outras condicionantes. O objectivo desta selecção é mostrar a forma de incorporar condicionalismos numa formulação matemática do problema e descrever uma rede neuronal de Hopfield que possibilite o desenvolvimento de uma heurística para o mesmo problema.

O P.E.H.E. descrito pode ser tomado como o problema da determinação de horários que respeitam os condicionalismos de 1º nível, isto é, constituem uma solução admissível e minimizam as violações relativas ao 2º nível de condicionalismos. Este problema pode ser formalizado em P.N.L.B. com os condicionalismos de 1º nível a originarem as restrições e em que os condicionalismos de 2º nível são assegurados pela função objectivo (Pato e Carrasco (1996)).

No modelo neuronal, com base em variáveis similares é definida uma função de energia a minimizar, composta por grupos de parcelas correspondentes quer aos condicionalismos de 1º

nível, quer aos de 2º nível. A diferenciação de importância destas condições faz-se no modelo neuronal recorrendo a ponderadores.

Nestes termos, consideram-se os seguintes parâmetros:

$s = (1, \dots, S)$ onde S é o número de salas disponíveis;

$d = (1, \dots, D)$ onde D é o número de disciplinas;

$t = (1, \dots, T)$ onde T é o número total de períodos de tempo por semana;

$p = (1, \dots, P)$ onde P é o número de professores;

$c = (1, \dots, C)$ onde C é o número de turmas;

M_d é o número máximo de aulas diárias para a disciplina d , para $d = (1, \dots, D)$;

N_d é o número de períodos de tempo semanais para a disciplina d previsto no plano curricular e na carga horária semanal do professor respectivo, para $d = (1, \dots, D)$;

$discP(p)$ é o conjunto de disciplinas leccionadas pelo professor p , para $p = (1, \dots, P)$;

$discC(c)$ é o conjunto de disciplinas assistidas pela turma c , para $c = (1, \dots, C)$;

$[B_{tt'}]$ é uma matriz quadrada de ordem T que exprime a penalização decorrente da ocupação de dois períodos de tempo (t e t');

$[C_{sd}]$ é uma matriz $S \times D$ tal que

$$C_{sd} = \begin{cases} 1 & \text{se a sala } s \text{ não é adequada à disciplina } d \\ 0 & \text{caso contrário} \end{cases}$$

$[L_{dd'}]$ é uma matriz quadrada tal que

$$L_{dd'} = \begin{cases} 1 & \text{caso as disciplinas } d \text{ e } d' \text{ sejam incompatíveis} \\ 0 & \text{caso contrário} \end{cases}$$

$[H_{pt}]$ é uma matriz $P \times T$ que regista a disponibilidade relativa ao professor p no período de tempo t ;

α , β e δ são os ponderadores associados às diferentes componentes da função objectivo do P.N.L.B.;

ρ_A , ρ_B , ρ_C , ρ_D , ρ_E e ρ_F são os ponderadores do modelo neuronal associados a cada grupo de restrições.

O processo de reformulação do P.N.L.B. num modelo baseado numa rede neuronal de Hopfield é semelhante ao proposto por Ramanujam e Sadayappan (1988). Proceder-se numa primeira etapa à identificação de cada variável binária do P.N.L.B.:

$$x_{sdt} = \begin{cases} 1 & \text{caso a sala } s \text{ fique ocupada pela disciplina } d \text{ no tempo } t \\ 0 & \text{caso contrário} \end{cases}$$

$$(s = (1, \dots, S); d = (1, \dots, D), t = (1, \dots, T))$$

A cada uma destas variáveis é associado um neurónio cujo estado, representado por x_{sdt} , pode variar no intervalo $[0,1]$ (neurónio activo, quando o estado ≈ 1 ; neurónio inactivo, quando o estado ≈ 0).

A Figura 2 ilustra genericamente os neurónios correspondentes à ocorrência ou não de cada disciplina d , no tempo t e na sala s .

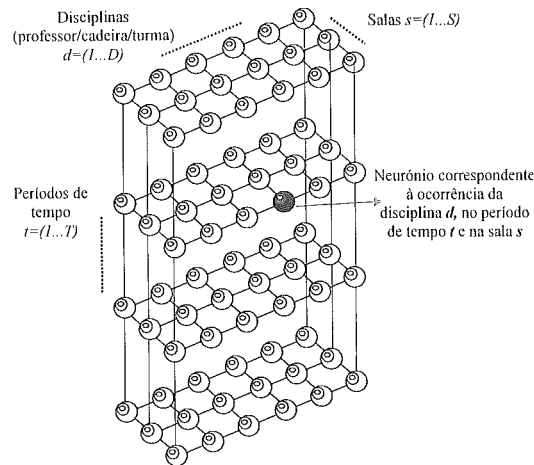


Figura 2 - Representação esquemática da rede neuronal

Na segunda etapa, sobre o conjunto de neurónios criados define-se uma função de energia do tipo da de Hopfield (3), cujo mínimo corresponde ao óptimo do P.N.L.B.. Assim, como se ilustra na Figura 3, cada condicionalismo inserido no modelo de P.N.L.B. (restrições e função objectivo) vai originar um grupo de parcelas da mesma natureza na função de energia e cuja anulação corresponde à satisfação do condicionalismo correspondente.

As restrições (5), que limitam a ocupação das salas ao máximo de uma por período de tempo, são integradas na função de energia através de um grupo (5') composto por $S \times T$ parcelas associadas a um ponderador (ρ_A). As restrições (6) impõem para cada professor um máximo de uma aula por período de tempo e são asseguradas na função energia por um grupo de $T \times P$ parcelas (6'). Para implementar as restrições (7), responsáveis pelo cumprimento do plano curricular previamente definido, são utilizadas D parcelas (7').

As restrições (8) são codificadas num grupo de T parcelas (8') o qual tem dois fins distintos: o primeiro é prevenir os conflitos horários resultantes da ocorrência simultânea de disciplinas comuns ao mesmo grupo de estudantes; o segundo objectivo é reforçar as restrições (6) evitando os eventuais conflitos horários para os professores. Esta redundância permite um incremento na qualidade das soluções obtidas no modelo neuronal relativamente à satisfação de (6), uma vez que tal preocupação se encontra distribuída por dois grupos distintos de parcelas (6') e (8').

A adequação das salas relativamente às disciplinas é assegurada por (9), assumindo na função a forma de T parcelas (9'). O último grupo de restrições (10), estabelece um limite máximo diário de aulas por disciplina e integra o modelo neuronal através de $3D$ parcelas (10'). Por motivo de simplificação este último grupo de restrições foi formulado partindo da instância exemplificativa apresentada na secção 5, na qual existem 3 períodos de tempo diários nos 3 dias disponíveis.

A função objectivo do P.N.L.B., expressões (11), (12) e (13), assume à partida uma natureza minimizadora sendo a sua codificação em termos neuronais perfeitamente directa. O somatório de todas as parcelas apresentadas define a função de energia.

Quando se procede à minimização desta função através de equações de actualização dos estados dos neurónios do tipo das Equações (4), assegura-se a convergência para um ponto de estabilidade, o qual pode corresponder a um mínimo local e, eventualmente, ao óptimo. A heurística neuronal cujo algoritmo será descrito na próxima secção destina-se a procurar um ponto que, após discretização dos estados dos neurónios, origine uma solução de qualidade do P.E.H.E., isto é, uma solução que viola "pouco" as restrições e corresponde a um "baixo" valor de z.

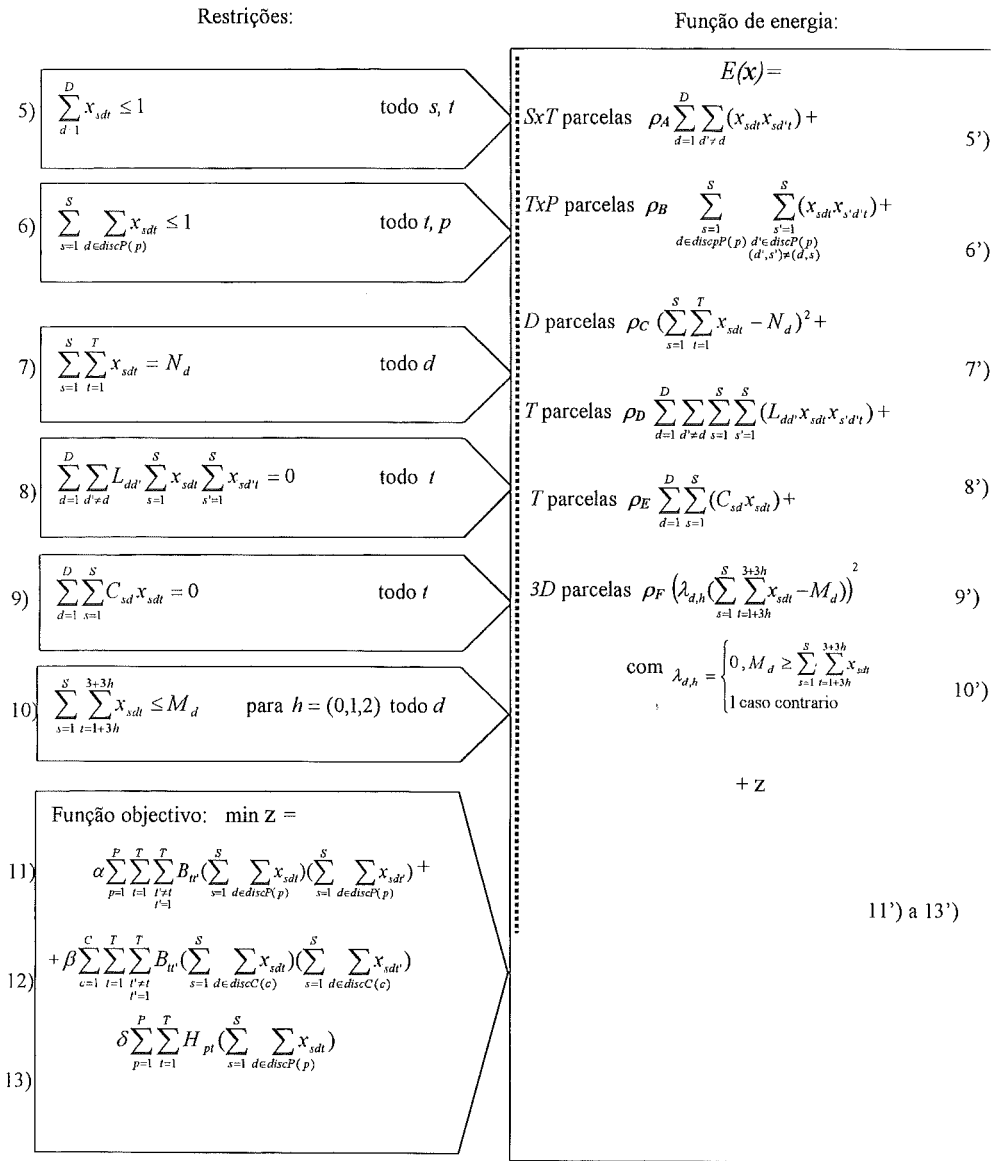


Figura 3 - Modelo de programação binária versus modelo neuronal

4. Descrição do Algoritmo

O algoritmo para a heurística neuronal, apresentado na Figura 4, pode ser decomposto em módulos.

No primeiro módulo são definidos os parâmetros e as variáveis utilizadas, é introduzido o conjunto de dados que caracteriza o problema (plano curricular, compatibilidade entre

disciplinas, disponibilidade dos professores, etc) e são ainda inicializados os ponderadores, o factor de ganho (G), o contador de iterações (*Iteração*), o número máximo de iterações (*MaxIter*), o valor da melhor solução discretizada (*Msol*) e o contador de pesquisas (*Npesq*).

No segundo módulo os estados dos neurónios da rede são inicializados com valores aleatórios adaptando o sistema utilizado por Wilson e Pawley (1988) para o Problema do Caixeiro Viajante, no qual os neurónios são configurados aleatoriamente até que a sua soma global seja aproximadamente igual ao número de neurónios activos pretendidos na solução final. Com este procedimento pretende-se acelerar o processo de convergência da heurística.

No terceiro procedimento, o estado de cada neurónio é individualmente actualizado de acordo com a equação apresentada anteriormente (Equação (4)). Optou-se por um esquema de actualização sequencial e síncrono, dada a sua maior rapidez e facilidade de implementação. Neste processo é definido um momento global de actualização no qual os neurónios procedem sequencialmente ao cálculo de U_i . Em seguida, o estado de cada neurónio é calculado através da aplicação da função sigmoideal, os neurónios com sinais de entrada negativos (com $U_i < 0$) deverão desactivar-se, aproximando o seu estado de 0; inversamente, os neurónios com sinais de entrada positivos (com $U_i > 0$) deverão activar-se, tendendo para 1. A sensibilidade de resposta em relação aos valores de entrada é determinada pelo factor de ganho.

No quarto módulo, são ajustados os ponderadores de acordo com o nível de satisfação do condicionalismo correspondente. Quando este se encontra fortemente "violado" ocorrerá um incremento no ponderador associado, obrigando a uma potencial satisfação dos seus requerimentos; inversamente, um condicionalismo verificado fará diminuir o ponderador respectivo. Exemplificando para o ponderador ρ_A , a actualização é efectuada do seguinte modo:

$$\rho_A = Pfix\rho_A + \frac{\rho_A(\text{parcelas } (5'))}{E(x)} Pvar\rho_A \quad (14)$$

em que, a constante $Pfix\rho_A$ define o valor mínimo assumido pelo ponderador ρ_A , enquanto que $Pvar\rho_A$ representa a componente variável do ponderador, atribuída de acordo com o peso na "violação total".

É ainda efectuado um incremento gradual do parâmetro G quando o algoritmo se aproxima do número máximo de iterações. Esta alteração destina-se a provocar uma convergência e diferenciação dos estados dos neurónios para valores próximos de 0 ou 1.

Após o incremento do contador de iterações testam-se os critérios de paragem do ciclo interno. A estabilização do valor da solução ($\Delta U = 0$) e o número máximo de iterações, foram os critérios para decidir sobre a realização de outra iteração. Quando este ciclo termina, consideram-se activos os neurónios com nível de activação (estado) superior a 0.5, os restantes são considerados inactivos. O valor da energia relativa à solução corrente discretizada (*Sol*) é comparado com o melhor até ao momento (*MSol*), sendo guardada essa solução caso seja menor.

Por último, o processo geral é reiniciado até que decorra o número de pesquisas definidas por $Npesq = 100$. No final, a solução heurística é dada pelos estados discretizados dos neurónios correspondentes ao valor da função de energia registado em $MSol$.

A implementação computacional do algoritmo foi desenvolvida em linguagem C (Borland C++), sendo executada num IBM-PC DX4 - 120MHZ. A performance registada pelo algoritmo nestas condições foi considerada satisfatória para a dimensão dos problemas testados, podendo ser muito superior quando vier a ser implementada num computador com capacidade de processamento em paralelo mais adequado à simulação de arquiteturas de computação distribuída.

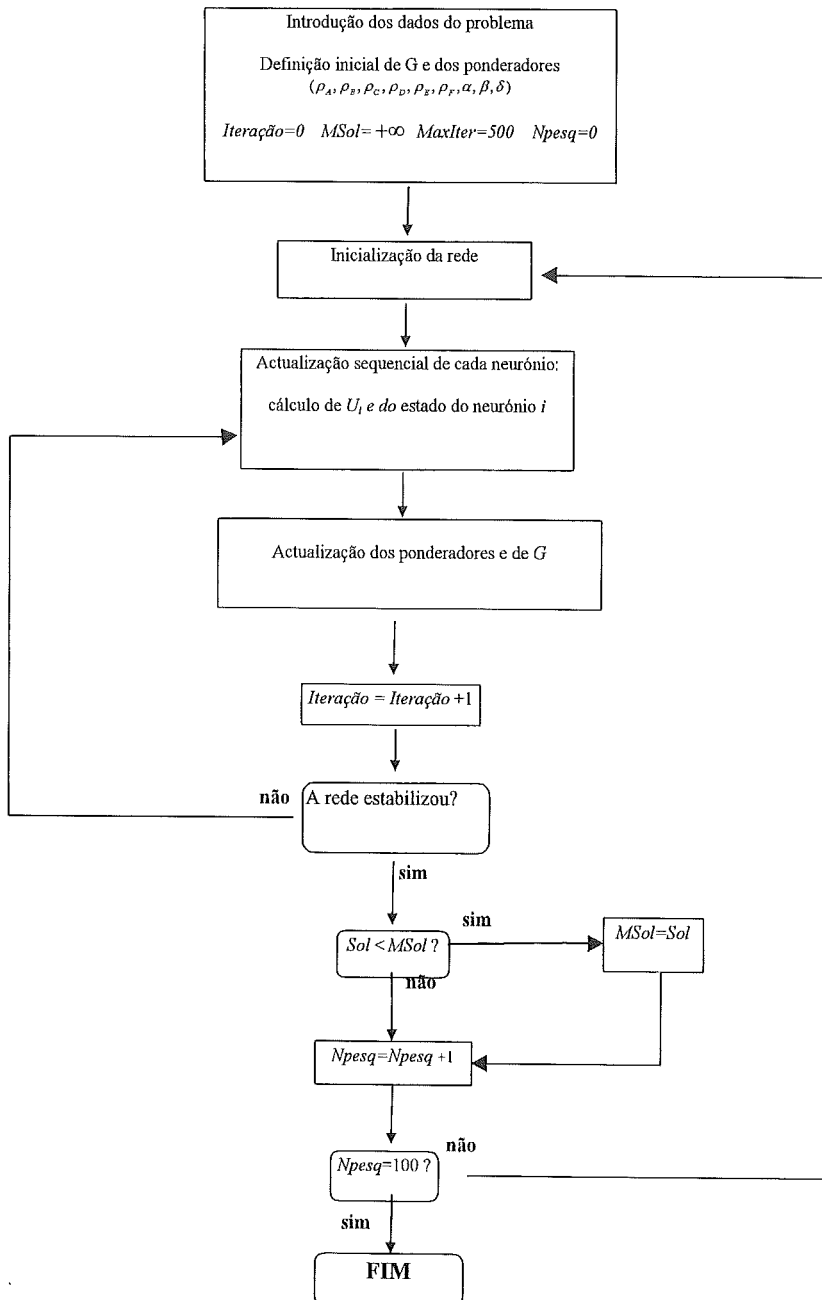


Figura 4 - Algoritmo neuronal

5. Aplicação a um P.E.H.E. exemplificativo e a instâncias pseudo-reais

Para ilustrar o funcionamento do algoritmo neuronal descrito anteriormente, apresenta-se uma instância exemplificativa do P.E.H.E. de dimensão reduzida, bem como quatro instâncias construídas para avaliação do comportamento da heurística face a problemas pseudo-reais com dificuldades distintas.

A instância exemplificativa apresentada encontra-se descrita em pormenor em Carrasco (1995) e representa a elaboração de horários numa instituição escolar hipotética com as seguintes características:

- um plano curricular com 25 aulas semanais;
- 3 dias de aulas, com 3 períodos de tempo diários, totalizando 9 tempos por semana;
- 6 professores, distribuídos pelas cadeiras de Matemática, Português, Geografia e Inglês;
- 4 turmas;
- 12 disciplinas;
- 3 salas (1 de grande capacidade, 1 de média capacidade e 1 de reduzida dimensão);
- índice de indisponibilidade dos professores de 33% (1 dia por professor);
- índice de incompatibilidade entre disciplinas na ordem dos 37%;
- índice de ocupação de salas elevado (93% de ocupação).

Definido o conjunto de dados iniciais, o procedimento heurístico passa pela constituição de uma rede neuronal com 324 neurónios (3 salas \times 12 disciplinas \times 9 tempos). Na resolução computacional foram testadas diversas combinações de parâmetros por forma a direccionar o algoritmo para a convergência para soluções consideradas satisfatórias, ou seja, que verifiquem os condicionalismos de 1º nível mesmo que não atendam à totalidade das preferências presentes nos condicionalismos de 2º nível.

Na prática a definição inicial destes valores revelou-se problemática, dada a dificuldade em equilibrar os ponderadores referentes aos condicionalismos de 1º nível - restrições - e os associados aos condicionalismos de 2º nível - função objectivo do P.N.L.B.. Constatou-se que, sem o teste de um conjunto adequado de parâmetros, o algoritmo tende a sobrevalorizar o aspecto ligado ao ponderador mais elevado, desprezando quase por completo os outros aspectos menos ponderados. Um equilíbrio aceitável dos parâmetros foi conseguido atribuindo uma ponderação significativamente reduzida aos condicionalismos de 2º nível, enquanto que aos de 1º nível foram atribuídas ponderações elevadas dado o seu carácter obrigatório. De salientar que estes ponderadores são modificados no decurso do algoritmo de acordo com o esquema de actualização já referido, com vista à satisfação integral das condições. Quanto ao factor de ganho G , foi testado com relativo sucesso o valor inicial de $G = 25$ proposto por Wilson e Pawley (1988) referente ao Problema do Caixeiro Viajante, verificando-se após diversas tentativas que o algoritmo regista um superior desempenho com o valor de $G = 50$.

Neste contexto, o algoritmo procedeu à determinação de soluções efectuando 100 pesquisas (cada uma podendo atingir o número máximo de 500 iterações), partindo de inicializações aleatórias distintas. O tempo total gasto no processo foi aproximadamente de 23 minutos. Este

tempo, por se tratar de um algoritmo neuronal simulado, não pode ser tomado como medida de eficiência comparativamente a outras abordagens. A melhor solução obtida surgiu durante a pesquisa nº 44 ($N_{pesq} = 44$), convergindo no final de 80 iterações. A evolução dos estados dos neurónios nessa pesquisa é graficamente resumida na Figura 5, a qual apresenta os estados dos neurónios no final das iterações 1 e 40. Nas figuras é ainda representada a solução final através da identificação com quadrados dos neurónios activos na solução final, tornando possível acompanhar as modificações nos estados até à sua estabilização final. A solução final satisfaz todos os condicionalismos de 1º nível considerados, o que implica a anulação de todas as parcelas da função energia correspondentes a estas restrições. Os resultados computacionais obtidos podem ser considerados bastante aceitáveis atendendo ao elevado índice de ocupação do problema (93%). Este facto, aliado a uma matriz de incompatibilidade entre disciplinas bastante densa (com 37% de conflitos, originada pela existência de disciplinas leccionadas pelos mesmos professores e optativas), tornava a instância pouco flexível e de resolução problemática. A existência por parte dos professores de uma indisponibilidade na ordem dos 33%, e o facto de algumas disciplinas necessitarem de salas específicas condiciona ainda mais esta situação, dificultando a pesquisa de uma solução sujeita a todas estas limitações.

Analisando a solução final traduzida nos horários do Quadro 1, verificamos que em termos de concentração das aulas foi atingido um resultado global muito satisfatório. Existem apenas dois furos, no horário da Turma B e do Professor 1. Quanto às disponibilidades evidenciadas previamente pelos professores, os resultados foram igualmente satisfatórios, obedecendo à maioria dessas preferências (96% estão satisfeitas). Existem apenas duas situações em que tal não se verifica (Professores 3 e 5, assinalados por (*) nos horários).

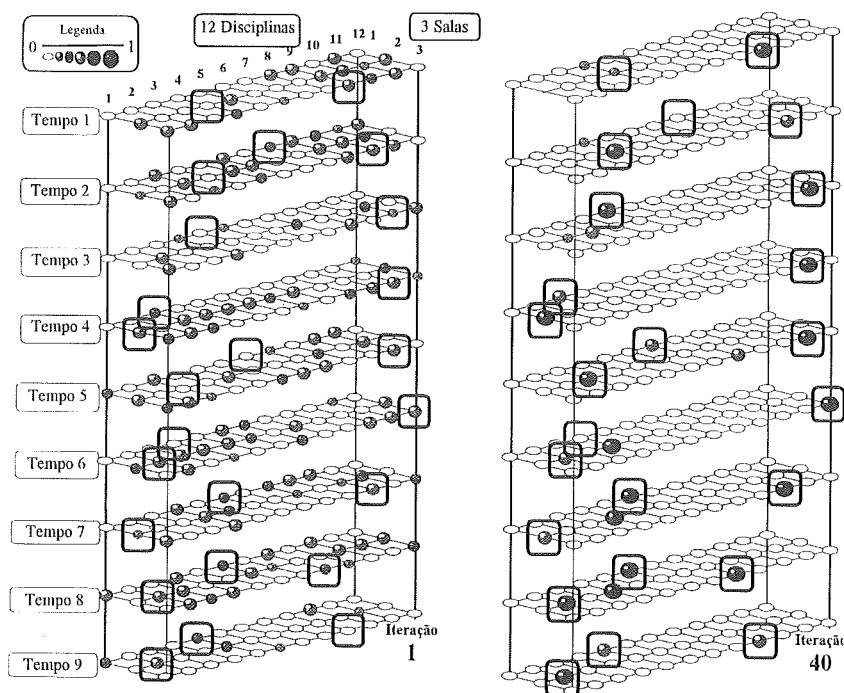


Figura 5 - Evolução da rede neuronal

Turma A

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12			Matemática	2	Matemática	2
13-16	Português	1			Português	3
16-19	Geografia	1	Inglês	3	Geografia	1

Turma B

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12	Português	3	Matemática	1	Geografia	1
13-16			Matemática	2	Geografia	1
16-19			Inglês	3	Português	3

Turma C

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12					Português	3
13-16	Português	3	Geografia	1	Matemática	2
16-19			Matemática	2	Matemática	2

Turma D

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12	Matemática	2	Português	3		
13-16	Matemática	2	Português	3		
16-19	Português	3	Matemática	1		

Professor 1

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12			Matemática-A	2	Matemática-A	2
13-16					Matemática-C	2
16-19			Matemática-C	2	Matemática-C	2

Professor 2

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12	Matemática-D	2	Matemática-B	1		
13-16	Matemática-D	2	Matemática-B	2		
16-19			Matemática-C	1		

Professor 3

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12					Geografia-B	1
13-16			Geografia-C	1	Geografia-B	1
16-19	Geografia-A(*)	1			Geografia-A	1

Professor 4

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12	Português-B	3				
13-16	Português-A	1			Português-A	3
16-19					Português-B	3

Professor 5

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12			Português-D	3	Português-C(*)	3
13-16	Português-C	3	Português-D	3		
16-19	Português-D	3				

Professor 6

Tempos	Dia 1	Sala	Dia 2	Sala	Dia 3	Sala
9-12						
13-16						
16-19			Inglês-A+B	3		

Quadro 1 - Horários para alunos e professores

Para estudar o comportamento do algoritmo proposto, face a problemas de horários com exigências e dimensões diferentes, procedeu-se à sua aplicação a um conjunto de instâncias de P.E.H.E. pseudo-reais de diferentes níveis de dificuldade, descritas em Carrasco (1995). A elaboração destes cenários, foi efectuada partindo de uma situação comum:

- uma semana escolar composta por 30 tempos horários de uma hora e meia de duração;
- 13 professores, repartidos pelas cadeiras de Matemática, Geografia, Português, Economia, Desenho, Inglês, Francês e Alemão;
- 9 turmas com igual dimensão, à excepção da 1 turma de dimensão superior, requerendo uma sala de maior capacidade;
- 25 disciplinas
- os professores consideram-se disponíveis em todos os períodos de tempo da semana.

Neste contexto, foi estabelecido um plano curricular comum a todos os cenários, no qual se define o número semanal de aulas para cada cadeira, bem como a turma e o docente envolvido. As turmas têm cerca de 17 aulas semanais, distribuídas por cadeiras nucleares (Matemática, Geografia, Português, Economia e Desenho) e opcionais (Inglês, Francês e Alemão). No total encontram-se programadas 120 aulas durante a semana. Refira-se o facto de algumas requererem salas especiais, como é o caso das opcionais (Inglês, Francês e Alemão), que devido ao elevado número de alunos reunidos necessitam de salas de grande capacidade. Definiu-se adicionalmente, por motivos pedagógicos, que nenhuma cadeira para a mesma turma deverá exceder o limite máximo de 2 aulas diárias.

Definidas as características comuns aos vários cenários, é agora apresentado o aspecto pelo qual estes diferem - o número de salas disponíveis. Optou-se por fazer a diferenciação através deste parâmetro, dado que ele condiciona profundamente a dificuldade do problema. Para o efeito consideraram-se quatro cenários distintos.

1º Cenário - 7 salas - Neste cenário, existem disponíveis 7 salas de aula, distribuídas pelos seguintes tipos: 4 salas de elevada capacidade e 3 de dimensões reduzidas. De acordo com os períodos de tempo definidos, o número potencial de aulas realizáveis é de 210 aulas (7 salas \times 30 tempos = 210 aulas), traduzindo-se num índice de ocupação de cerca de 57% ($\frac{120 \text{ aulas programadas}}{210 \text{ aulas possíveis}} = 57.1\%$). Este valor indicativo não considera ainda as limitações adicionais decorrentes do facto das disciplinas necessitarem de salas adequadas, resultando num índice de combinações incompatíveis sala/disciplina de 23%.

2º Cenário - 6 salas - Este cenário, caracteriza-se por ter 6 salas de aula, tendo sido eliminada uma sala de grande capacidade. A dificuldade do problema é acrescida dada a diminuição do número potencial de aulas realizáveis (6 salas \times 30 tempos = 180 aulas). O índice de ocupação aumenta para 66,7%. O plano das salas é idêntico ao anterior removida uma sala, provocando um aumento do índice de combinações incompatíveis sala/disciplina para 29%.

3º Cenário - 5 salas - Dispondo nesta situação de um máximo de 5 salas, através da eliminação de mais uma sala de dimensões elevadas vai tornar-se ainda mais difícil a determinação de uma solução admissível para o P.E.H.E., facto comprovável através do índice de ocupação que atinge os 80%. O índice de combinações incompatíveis sala/disciplina sobe para 34%.

4º Cenário - 4 salas - Neste último cenário, no qual existem apenas 4 salas de aula (uma única de elevada capacidade), o índice de ocupação atinge os 100%, correspondendo a um problema saturado. Nesta situação todas as salas deverão ser permanentemente ocupadas. O índice de combinações incompatíveis sala/disciplina atinge os 43%. Por estas razões o 4º cenário é o mais problemático, existindo maior dificuldade em determinar uma solução que verifique simultaneamente todas as restrições - condicionalismos de 1º nível.

Na resolução computacional foram utilizadas redes neuronais com 5250, 4500, 3750 e 3000 neurónios, respectivamente para o 1º, 2º, 3º e 4º cenários considerados. Em termos gerais o processo de implementação efectuou-se num contexto idêntico ao proposto para a instância exemplificativa. A definição dos parâmetros utilizados pelo algoritmo foi difícil, obrigando a um processo demorado de experimentação. Foi possível constatar que os parâmetros

necessitam de sofrer constantes ajustamentos para acompanhar as modificações nos dados do problema. Estabeleceu-se um factor de ganho mais baixo comparativamente ao definido no exemplo, para tornar o algoritmo mais "sensível" na pesquisa de soluções. Esta modificação traduz-se também num aumento do tempo computacional exigido, dado que os neurónios demoram mais tempo a definir o seu estado final. O esquema de actualização utilizado foi novamente síncrono e sequencial. Foram igualmente experimentadas outras formas de actualização alternativas (assíncrona, não sequencial), não evidenciando melhorias significativas na qualidade das soluções obtidas.

Após a execução do algoritmo para os diferentes cenários obtiveram-se os resultados que se encontram resumidos no Quadro 2.

Problemas	Índice de ocupação	Índice de combinações incompat.	Restrições (condicionalismos de 1º nível)						Número de furos	
			5	6	7	8	9	10	Professores	Alunos
1º cenário	57%	23%	0	0	0	0	0	0	28	40
2º cenário	67%	29%	0	0	0	0	0	0	31	35
3º cenário	80%	34%	0	0	0	0	24	0	27	30
4º cenário	100%	43%	0	0	2	0	35	0	25	36

Quadro 2 - Resultados computacionais

À esquerda, apresentam-se dois índices caracterizadores da dificuldade de cada cenário, o índice de ocupação e de combinações incompatíveis sala/disciplina. Em seguida, no centro da tabela, apresentam-se para cada cenário, o número total de violações por cada tipo de restrição. À direita apresenta-se uma medida relativa à qualidade dos horários obtidos: o número total de furos existentes em cada grupo de horários. Este valor é representativo do grau de satisfação da função objectivo do P.N.L.B..

Como se pode verificar, a heurística apenas em duas instâncias (3º e 4º cenários) não forneceu uma solução admissível, mas os condicionalismos de 1º nível foram quase integralmente respeitados, o que é muito satisfatório se atendermos, por exemplo, à dificuldade existente na resolução do 4º cenário que corresponde a uma situação de saturação. Neste caso só foi possível encontrar uma solução à custa da violação de algumas restrições, principalmente da restrição de adequação das salas (9).

6. Conclusões

Neste artigo descreveu-se sinteticamente uma formalização em P.N.L.B. para o P.E.H.E. e uma heurística suportada por uma rede neuronal de Hopfield. Apresentam-se ainda resultados da sua aplicação a uma instância exemplificativa e a diferentes instâncias pseudo-reais.

Concluiu-se que a referida rede neuronal de Hopfield permite uma modelização directa das variáveis do P.N.L.B.. Existe assim uma transparência formal entre o estado dos neurónios e o valor das respectivas variáveis binárias, o que facilita a observação da evolução do algoritmo neuronal até à solução final.

Contudo, o processo de elaboração de uma rede neuronal específica para o P.E.H.E. proposto revelou-se pouco flexível, sendo difícil a posterior introdução de modificações que obrigariam a uma total reformulação da função de energia do modelo neuronal. No que respeita à determinação de valores para os parâmetros deste modelo que garantam a convergência do algoritmo para soluções admissíveis e de qualidade é de salientar a considerável dificuldade sentida. Para cada instância do problema os parâmetros são encontrados através de um processo de experimentação com vista a orientar a convergência. Estes dois factos são bastante limitativos do ponto de vista prático se considerarmos a flexibilidade naturalmente exigida para um método destinado à futura abordagem de problemas reais.

Ao nível das soluções obtidas por aplicação da heurística neuronal verificaram-se dois aspectos distintos.

Em relação aos diversos condicionalismos de 1º nível para a elaboração de horários a abordagem neuronal mostrou-se bastante eficiente na sua verificação, permitindo encontrar, em todos os casos estudados, soluções satisfatórias em termos do cumprimento das restrições.

No que respeita à qualidade da solução (condicionalismos de 2º nível), ou seja, ao cumprimento das orientações expressas pela função objectivo do P.N.L.B., a heurística neuronal demonstrou algumas deficiências. Tal facto deve-se à forma como é incluído na função energia todo o conjunto exigências do problema, dificultando o equilíbrio de importância entre os condicionalismos de 2º nível e os de 1º nível. Na prática o algoritmo tende a sobrevalorizar um destes grupos.

Neste contexto, verificamos que a heurística baseada no modelo neuronal, embora necessite de algum estudo adicional sobre o processo de parametrização, se mostrou adequada para problemas nos quais, tal como no P.E.H.E., o objectivo não é muito rígido e as restrições são maleáveis, podendo eventualmente ser violadas de acordo com a sua importância relativa. Genericamente a abordagem neuronal demonstra algumas potencialidades inovadoras, não permitindo de momento a comparação directa dos tempos de execução com os métodos clássicos de resolução, já que a sua implementação computacional é vocacionada para o processamento distribuído.

Referências

- [1] Braun, M., *Differential equations and their applications*, Springer-Verlag, New York (1975).
- [2] Carrasco, M. P., *Redes neuronais na elaboração de horários escolares*, Dissertação de Mestrado, Departamento de Matemática, Instituto Superior de Economia e Gestão - U.T.L. (1995).
- [3] Carter, M. W., "Timetabling literature", disponível em <ftp://ftp.cs.nott.ac.uk/ftp> (1995).
- [4] Gislén, L., Peterson, C., Söderberg, B., "Complex Scheduling with Potts Neural Networks", *Neural Computation* 4 (1992) 806-831.
- [5] Hopfield, J., "Neural networks and physical systems with emergent collective computational ability", *Proceedings of the National Academy of Sciences* 79 (1982) 2554-2558.
- [6] Hopfield, J., "Neurons with graded response have collective computational properties like those two-state neurons", *Proceedings of the National Academy of Sciences* 81 (1984) 3088-3092.
- [7] Hopfield, J., Tank, D., "Neural computation of decisions in optimization problems", *Biological Cybernetics* 52 (1985) 141-152.
- [8] McCulloch, W., Pitts, W., "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics* 5 (1943) 115-133.

- [9] Pato, M. V., Carrasco, M. P., "A binary nonlinear formulation for a secondary school timetabling problem", Working paper nº 7/96, Centro de Investigação Operacional, Faculdade de Ciências - U.L. (1996).
- [10] Ramanujam, J., Sadayappan, P., "Optimization by neural networks", *Proceedings of the I.E.E.E. Second International Neural Networks Conference 2* (1988) 325-332.
- [11] Sharda, R., "Neural networks for the MS/OR analyst: an application bibliography", *Interfaces* 24 (1994) 116-130.
- [12] Werra, D. de, "An introduction to timetabling", *European Journal of Operational Research* 19 (1985) 151-162.
- [13] Wilson, G., Pawley, G., "On the stability of the Travelling Salesman Problem algorithm of Hopfield and Tank", *Biological Cybernetics* 58 (1988) 63-70.
- [14] Yu, T., "Timetable scheduling using network algorithms", *Proceedings of the I.E.E.E. International Joint Conference on Neural Networks* 1 (1990) 270-284.

UM MÉTODO DE PROGRAMAÇÃO LINEAR MULTI-OBJECTIVO BASEADO NA PROJEÇÃO DOS GRADIENTES DAS FUNÇÕES OBJECTIVO

João Paulo Costa

João Carlos Clímaco

Faculdade de Economia da Univ.Coimbra
INESC - Coimbra

Abstract

This paper presents an interactive method for decision support to multiobjective linear programming problems. This method is based on the projection of the objective function gradients over the non-dominated region of the problem. It allows an oriented and continuous search for the preferred solution. The method uses also the dual variables, according to the achievement scalarizing program used in the projections, to characterize the non-dominated hiper-faces of the problem.

The most representative multiobjective linear programming methods and systems, based on reference points, are also reviewed in this paper. The purpose of this revision is to transmit a global view of the evolution of these methods and systems, allowing their comparison with the new presented method.

Finally the method's usage is illustrated through a small example.

Resumo

Neste trabalho apresenta-se um método interactivo de apoio à decisão em problemas de programação linear multiobjectivo. O método baseia-se na projecção de pontos de referência e dos gradientes das funções objectivo sobre a região não dominada, permitindo uma procura orientada e contínua da solução preferida. O método serve-se, ainda, das variáveis duais, segundo o programa escalarizante de realização utilizado nas projecções, para caracterizar as diferentes hiperfaces que compõem uma região não dominada.

Neste artigo referem-se também, de forma breve, os mais representativos métodos e sistemas interactivos de programação linear multiobjectivo baseados em pontos de referência, com o intuito de fornecer ao leitor uma visão integrada da área e permitindo o confronto da evolução destes métodos com o que aqui se propõe.

Finalmente ilustra-se a utilização do método com um pequeno exemplo.

Keywords

Multiobjective linear programming, reference points, oriented continuous search.

1. Introdução

Utilizando gráficos interactivos e projecções de pontos de referência a variar segundo uma direcção, Korhonen desenvolveu uma abordagem interactiva para os problemas de programação linear multiobjectivo. Esta abordagem resultou numa ferramenta computacional (VIG-Korhonen (1987)), pioneira em apoiar uma pesquisa da região não dominada segundo uma determinada direcção.

Embora semelhante em alguns pontos, o método que se apresenta neste artigo, e que se designará por método C^2 , baseando-se na projecção dos gradientes das funções objectivo, difere do VIG no que se relaciona com:

- a) A escolha da direcção de pesquisa. Este novo método permite ao agente de decisão escolher a direcção de pesquisa por forma a melhorar simultânea e explicitamente o valor de várias funções objectivo, sabendo quais as funções cujo valor terá forçosamente que degradar. Para além disso, sabe-se ainda qual a percentagem relativa de variação de cada valor (melhorando ou piorando) e as possibilidades da sub-região não dominada em pesquisa. O agente de decisão tem portanto os instrumentos necessários para melhorar o controlo da direcção de pesquisa.
- b) A determinação da sub-região não dominada que oferece as perspectivas mais interessantes para uma pesquisa local. O método descrito neste artigo caracteriza, a partir de informação dual, a sub-região (hiperface) não dominada actualmente sob pesquisa, apresentando ainda a caracterização das regiões não dominadas vizinhas. Com esta informação o agente de decisão pode mudar de sub-região para continuar a sua pesquisa, sem ter que viajar continuamente segundo uma determinada direcção.
- c) A estrutura das preferências. Dado o mecanismo de cálculo, o agente de decisão pode estruturar as suas preferências em três níveis diferentes: (i) o espaço de decisão (neste nível o controlo exercido é inferior ao possível nos outros níveis) e o espaço dos critérios, utilizando num nível (ii) o valor das funções objectivo e noutra (iii) os pontos de referência. Os instrumentos permitem assim lidar com diferentes tipos de situação e enquadramento racional.

Este artigo inicia-se apresentando um resumo da evolução histórica dos métodos de ponto de referência, salientando-se os pontos relevantes para entender as razões do desenvolvimento do novo método. Pretende-se assim descrever a evolução de alguns conceitos teóricos e das técnicas que os implementam, culminando no sistema de apoio à decisão VIG. Refere-se ainda alguma da investigação posterior com vista a melhorar este sistema. Na secção seguinte, a secção 3, apresenta-se detalhadamente o método objecto desta comunicação. Finalmente na secção 4 exemplifica-se o método através do tratamento de um problema de pequenas dimensões.

2. Métodos de Ponto de Referência

O método de STEM, Benayoun et al (1971), foi o primeiro procedimento interactivo a ter impacto no campo da programação linear multiobjectivo, veja-se Steuer (1986). A sua abordagem é a da redução da região admissível. Em cada iteração o método apresenta uma solução não dominada¹, minimizando a distância de Tchebycheff à solução ideal, e o agente de decisão é chamado a especificar as quantidades que está disposto a sacrificar nas funções objectivo cujo valor considera satisfatório, permitindo, assim, melhorar o valor das restantes.

¹ Na sua versão original este método não distinguia entre soluções eficientes e fracamente eficientes.

Esta especificação traduz-se em restrições lineares, obrigando as funções objectivo que o agente de decisão considerou de valor satisfatório a serem superiores² ao seu valor actual menos a quantidade que está disposto a sacrificar. Na iteração seguinte recalcula-se a solução não dominada com a região admissível reduzida, face a estas novas restrições, e com a ponderação da métrica de Tchebycheff alterada com muito menos peso para as funções objectivo correspondentes às novas restrições. A procura da solução não dominada satisfatória evolui, assim, aos 'saltos' sobre zonas da região não dominada cada vez menores, sendo a ferramenta de controlo, para a obtenção da solução não dominada seguinte, algo rudimentar, pois não permite grande 'sensibilidade' na variação dos valores das funções objectivo.

Vincke (1976) apresentou uma variante do método de STEM que utiliza a análise de sensibilidade clássica aos termos independentes das restrições do programa escalarizante de realização com o intuito de melhorar a interactividade do método permitindo que os 'saltos', acima referidos, sejam mais suaves. Embora não tenha sido completamente explorada, esta variante do método de STEM propôs alguns princípios de interactividade que mais tarde com o método de PARETO RACE, Korhonen e Wallenius (1988), deram os seus frutos, mas com um enquadramento teórico proposto por Korhonen e Laakso (1984), tal como se refere alguns parágrafos à frente.

Um outro método, apresentado por Steuer e Choo (1983), usa a distância ponderada e aumentada de Tchebycheff. Este método segue uma estratégia de concentrar progressivamente uma amostra de soluções não dominadas, a partir das preferências manifestadas pelo agente de decisão ao escolher a solução preferida da amostra anterior. Numa primeira iteração, depois de obter a solução ideal, forma-se um grupo de vectores dispersos. Estes vectores definem um grupo de distâncias ponderadas e aumentadas de Tchebycheff, dispersas de forma a obter-se uma amostra de todo o conjunto de soluções não dominadas através da minimização destas distâncias. Então, o agente de decisão selecciona uma solução como sendo a sua preferida de todo o conjunto. Nas iterações seguintes (tantas quantas um número fixo e pré-definido, a menos da interrupção do processo por o agente de decisão se encontrar satisfeito) volta-se a formar um conjunto de vectores idêntico aos anteriores, mas cada vez mais concentrado em redor do vector que conduziu à solução não dominada preferida pelo agente de decisão na iteração anterior. As soluções associadas a cada conjunto de vectores são determinadas e o agente de decisão, em cada iteração, escolhe aquela que prefere e que servirá de ponto de partida à próxima iteração.

O método 'Satisficing Trade-off', apresentado por Nakayama e Sawagari (1984), é semelhante ao método de STEM, no entanto o agente de decisão para além de indicar quais os critérios que não se importa de degradar, indica também explicitamente quais os critérios a melhorar. Por outro lado, os níveis de aspiração do agente de decisão são utilizados directamente para a obtenção dos pesos da distância pesada de Tchebycheff ao ponto ideal.

² Se a função objectivo for a maximizar.

Neste método tenta-se encontrar uma solução eficiente³ que seja satisfatória, isto é, cujos valores nos diferentes critérios sejam superiores ou iguais aos valores indicados pelo agente de decisão como níveis de aspiração. O procedimento inicia-se pelo cálculo da solução ideal. Seguidamente o agente de decisão fixa um primeiro conjunto de níveis de aspiração. A partir da solução ideal e dos níveis de aspiração determinam-se os pesos de uma distância pesada de Tchebycheff ao ponto ideal, que quando minimizada permite obter uma solução fracamente eficiente. Baseado nos valores das funções objectivo desta solução o agente de decisão classifica os critérios em três classes: i) a classe dos critérios que pretende melhorar, ii) a dos que concorda em relaxar e iii) a classe dos critérios cujo valor aceita. Para as duas primeiras classes são pedidos ao agente de decisão novos níveis de aspiração, para a terceira os níveis de aspiração passam a ser iguais aos valores das funções objectivo. Seguidamente aplica-se um procedimento baseado em multiplicadores Lagrangeanos, que verifica se os novos níveis de aspiração constituem um ponto admissível (alcançável - dado que se trata do espaço dos objectivos). Se este for admissível volta-se a minimizar a distância pesada de Tchebycheff ao ponto ideal, mas agora com pesos decorrentes dos novos níveis de aspiração, por forma a determinar a solução fracamente eficiente que lhe corresponde. Repare-se que esta solução fracamente eficiente pode apresentar, para os critérios, valores melhores que os apontados pelos níveis de aspiração. Caso o ponto não seja admissível, pode pedir-se ao agente de decisão novos níveis de aspiração ou voltar a minimizar a distância de Tchebycheff correspondente e reiniciar o processo, dependendo da dificuldade experimentada pelo agente de decisão em encontrar níveis de aspiração que constituam um ponto admissível.

Sob a forma de sistema de apoio à decisão surge o sistema DIDAS - Dynamic Interactive Decision Analysis & Support, desenvolvido por Lewandowski et al (1988) e baseado também no princípio de optimização por pontos de referência. O sistema DIDAS utiliza uma função escalarizante de realização do ponto de referência sobre a região não dominada. Esta função difere das utilizadas na programação por metas ('goal programming') e nos métodos referenciados anteriormente (funções baseadas em normas), por ser coerente na ordem. Uma função coerente na ordem mantém-se monótona mesmo quando o ponto de referência atravessa a fronteira não dominada. Simplificadamente pode dizer-se que uma função coerente na ordem muda automaticamente da minimização de uma norma para a maximização dessa norma (ou outra, desde que adequada), quando os níveis de aspiração se tornam factíveis (ponto de referência admissível). As funções coerentes na ordem podem dividir-se em funções que representam uma ordem e em funções que a aproximam. Distinguem-se entre si por as primeiras poderem originar resultados fracamente eficientes e as segundas garantirem resultados não dominados. A função utilizada pelo sistema DIDAS aproxima uma ordem através da perturbação de cada função objectivo pelo conjunto das restantes. Esta perturbação é controlada por um parâmetro, ϵ , que representa a capacidade de discriminação do agente de decisão. Isto é,

³ Veja-se nota 1.

há soluções eficientes que, para um determinado agente de decisão, não se distinguem das fracamente eficientes pois diferem muito pouco de uma outra solução eficiente. Assim, o sistema DIDAS apenas apresenta soluções propriamente⁴ eficientes com pelo menos uma razão melhoria/degradação dos valores das funções objectivo limitadas por $1/\epsilon$ ou ϵ de acordo com o numerador e denominador da razão. No DIDAS o agente de decisão controla as soluções eficientes, calculadas pelo sistema, modificando os seus níveis de aspiração. Para além disso, o agente de decisão pode realizar uma operação de análise denominada por 'varrimento multidimensional'. Esta operação, mais próxima da análise de cenários que da de sensibilidade, consiste em seleccionar um ponto não dominado, aceitá-lo como base (âncora) e calcular outros pontos não dominados, fazendo variar o ponto de referência por incrementos, dentro dos limites do ponto utópico e do ponto nadir. No que respeita à convergência para a solução preferida Lewandowski et al (1988) propõem a utilização de um 'varrimento direccionado', utilizando agora duas âncoras e fazendo variar o ponto de referência sobre um vector entre elas. Se se modificarem as âncoras por forma a variar as direcções de procura, então, eventualmente, o sistema converge para a solução preferida, veja-se Korhonen e Laakso (1986).

Um outro sistema de apoio à decisão, o VIG - Visual Interactive Goal Programming, ainda hoje disponível comercialmente, foi proposto por Korhonen (1987). Note-se que este sistema mudou ocasionalmente de nome, veja-se Korhonen e Wallenius (1992), para VIG - Visual Interactive Graphic support. Com este sistema pode procurar-se no conjunto de soluções eficientes⁵ a solução preferida pelo agente de decisão, privilegiando uma função objectivo em relação às outras e definindo o passo que mede a proximidade das soluções que se obtêm em cada interacção. Uma semi-recta do espaço dos critérios, com início no ponto de referência determinado pelos níveis de aspiração, é projectada iterativamente sobre a região eficiente, gerando uma trajectória que pode ser percorrida num processo de busca. O sistema VIG torna este processo dinâmico através da apresentação dos valores das funções objectivo num gráfico de barras cuja dimensão varia à medida que o agente de decisão indica qual a função que pretende melhorar (repare-se que em cada interacção só pode fornecer esta indicação sobre uma função objectivo, podendo perder as indicações dadas anteriormente) e a proximidade da próxima solução (passo). O mecanismo de cálculo do sistema VIG é baseado na análise de sensibilidade clássica aos termos independentes das restrições do problema escalarizante de realização associadas às funções objectivo. Trata-se de um processo semelhante ao utilizado por Vincke (1976) na variante ao método de STEM referida acima.

Os fundamentos do sistema de apoio à decisão VIG foram modificados por Halme (1992) através da introdução de melhorias nos instrumentos de controlo da direcção de pesquisa,

⁴ O conceito de solução propriamente eficiente só tem relevância em programação linear neste contexto das funções que aproximam uma ordem. Iserman (1974) demonstrou que todas as soluções eficientes em programação linear são também propriamente eficientes, desde que o limite da razão melhoria/degradação (isto é, ϵ) seja suficientemente pequeno.

⁵ A proposta inicial podia resultar na obtenção de soluções fracamente eficientes, no entanto veio a ser modificada de acordo com a proposta de aproximação de ordem de Lewandowski.

fundamentando-se no trabalho de Halme e Korhonen (1989). Neste trabalho subsistem, no entanto, alguns problemas de cálculo no caso de soluções degeneradas. A modificação introduzida prende-se com a direcção de referência, isto é, a direcção de deslocação do ponto inicial. No VIG, ao projectar-se a direcção de referência sobre a região não dominada obtém-se a trajectória de pesquisa das soluções não dominadas. Inicialmente Korhonen e Laakso (1984) propuseram a utilização de níveis de aspiração: depois de encontrar uma primeira solução não dominada, pede-se ao agente de decisão para fornecer novos níveis de aspiração, sendo a direcção de referência calculada através da diferença entre o ponto de referência associado aos níveis de aspiração e a solução não dominada (no espaço dos critérios). O VIG foi implementado de tal forma que a direcção de referência é especificada dinamicamente pelo próprio sistema, a partir de unidades de escala e do desejo do agente de decisão melhorar uma função objectivo. Em nenhum destes casos a direcção de referência é necessariamente admissível. Halme e Korhonen (1989) propõem uma modificação em que a direcção de referência é admissível e não dominada, continuando, no entanto, a ser explicitada no espaço dos critérios. O processo de escolha da direcção de referência, nesta proposta, passa pela obtenção prévia de um conjunto de vectores a partir do quadro do método simplex associado à solução óptima do problema escalarizante de realização. É aqui que há problemas não resolvidos caso esta solução seja degenerada no sentido do problema escalarizante de realização. Os vectores a obter são nem mais nem menos do que as arestas de intersecção das restrições auxiliares do problema escalarizante, associadas às funções objectivo, com o hiperplano não dominado em pesquisa. Se a solução não dominada actual estiver sobre um hiperplano constituído pela intersecção de duas ou mais restrições do problema original (conduzindo a uma solução degenerada para o problema escalarizante, embora possa nem ser básica para o problema original), os vectores a obter não estão sobre a mesma restrição do problema original. Ora, nestas condições, o cone gerado pelos vectores inclui direcções dominadas, pelo que esta técnica é de aplicação muito limitada. Isto é, quando se chega a uma aresta, durante a pesquisa, a técnica deixa de ser aplicável. No entanto, esta tentativa de melhorar a obtenção e o controlo da direcção de referência demonstra a necessidade, que os próprios investigadores envolvidos no sistema VIG sentem, de fornecer mais apoio ao agente de decisão para a determinação da direcção de deslocamento. Após a obtenção do conjunto de vectores, é pedido ao agente de decisão que especifique uma direcção de pesquisa dentro do cone gerado pelo conjunto destes vectores. Para além da determinação da direcção de referência há ainda a referir que Korhonen e Halme (1989) propõem ainda um conjunto de condições matemáticas como critérios de paragem da pesquisa. Este critério é, no entanto, de aplicação limitada, pois só se pode utilizar em problemas onde se assuma a existência, mesmo que implícita, de uma função valor pseudo-côncava.

Na secção seguinte apresenta-se um novo método que se integra nesta área dos métodos de ponto de referência e que apresenta vantagens face a estes, tal como é indicado na secção 1.

3. O Novo Método

O método inicia-se pela projecção do ponto de referência, definido pelo agente de decisão, sobre a fronteira eficiente (da imagem da região admissível, no espaço dos critérios), obtendo-se uma solução não dominada inicial. A secção 3.1 apresenta o programa escalarizante de realização utilizado para o efeito. A característica dual da hiperface ou hiperfaces não dominadas sobre as quais a solução não dominada se encontra é então apresentada ao agente de decisão. Esta característica dual é obtida de acordo com a secção 3.6. Seguidamente o agente de decisão deve pronunciar-se sobre se quer pesquisar uma determinada sub-região não dominada ou se prefere procurar noutra vizinha (veja-se secção 3.7 para as hiperfaces vizinhas). Considere-se que o agente de decisão escolheu uma sub-região não dominada. Para mover a projecção do ponto de referência sobre esta sub-região é necessário encontrar uma direcção de pesquisa. Para isso projectam-se os gradientes das funções objectivo sobre esta sub-região, veja-se secção 3.2, definindo-se p direcções⁶ (sendo p o número de funções objectivo). O conjunto das p direcções permite viajar sobre a sub-região não dominada sob pesquisa. O produto interno do gradiente normalizado de cada função objectivo com a sua projecção normalizada define um valor h_k ($k = 1, \dots, p$), que mede a proximidade máxima ao critério k que se pode conseguir quando se viaja sobre a sub-região. O agente de decisão ciente destes valores define interactivamente, para cada função objectivo, a proximidade mínima que considera aceitável. A secção 3.3 explicita este processo. Os cálculos necessários à exploração de uma sub-região não dominada segundo uma direcção, são apresentados na secção 3.4. Note-se que estes cálculos se realizam no espaço de decisão. Se durante a deslocação da solução não dominada se atingir uma intersecção com outra, ou outras, restrições (que passam também a ser activas) e se o agente de decisão pretender mudar para essa nova hiperface associada às novas restrições, ele deve primeiro obter a sua característica dual (verificando-se se a face é ou não eficiente), voltando seguidamente a encetar o processo de escolha da direcção de pesquisa. Se por outro lado, o agente de decisão pretender mudar de direcção sem sair da mesma face, pode fazê-lo sem recalculer a característica dual da face em que se encontra.

Para cada solução não dominada actual pode determinar-se o ponto de referência que lhe daria origem considerando-se as unidades de escala iniciais. Pode também encontrar-se a direcção de movimento dos pontos de referência, equivalente à direcção de movimento dos pontos não dominados. Veja-se a secção 3.5. Assim permite-se ao agente de decisão tomar opções ao nível das suas aspirações, utilizando os pontos de referência correspondentes aos pontos não dominados que vai encontrando.

Os fundamentos do método que aqui se propõe surgem em Costa e Clímaco (1994).

⁶ Algumas destas direcções podem ser coincidentes. Em casos limite podem também resultar em pontos, não sendo possível a deslocação (caso, por exemplo, em que a hiperface não dominada é constituída por um ponto).

3.1 Programa linear escalarizante de realização

O problema linear multiobjectivo, objecto deste estudo, pode ser escrito da seguinte forma, sem perda de generalidade:

$$\max z_1 = c^1 x; \max z_2 = c^2 x; \dots; \max z_p = c^p x$$

sujeito a:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i, \quad 1 \leq i \leq m$$

$$x_j \geq 0, \quad j = 1, \dots, n$$

onde p é o número de funções objectivo e $c^k \in \mathbb{R}^n$ é o gradiente da função objectivo k . As restrições de não negatividade corresponderão, na exposição seguinte, às restrições cujo índice é maior ou igual a $m+1$ e menor ou igual a $m+n$. Seja $q^* \in \mathbb{R}^p$ o ponto de referência no espaço dos critérios. Este ponto pode ou não ser admissível. A função escalarizante de realização utilizada, $s(q^*, z)$, é uma versão simplificada da de Lewandowski e Wierzbicki (1988):

$$s(q^*, z) = \min_{1 \leq k \leq p} \frac{z_k - q_k^*}{s_k} + \frac{\varepsilon}{p} \sum_{k=1}^p \frac{z_k - q_k^*}{s_k}$$

O parâmetro ε deve ser positivo, mesmo se muito pequeno. Se este parâmetro for igual a zero, então o máximo da função $s(q^*, z)$ pode corresponder a soluções fracamente eficientes. s_k é a unidade de escala do critério k . Se se maximizar a função $s(q^*, z)$ sujeita à região admissível do problema original obtém-se como resultado um ponto não dominado. A melhor forma de resolver este problema é utilizando um programa linear auxiliar que se denominará programa linear escalarizante de realização:

$$\max \left(\alpha + \frac{\varepsilon}{p} \sum_{k=1}^p \frac{z_k - q_k^*}{s_k} \right)$$

$$\text{s.a: } c^k x = z_k \quad 1 \leq k \leq p$$

$$\alpha \leq \frac{z_k - q_k^*}{s_k} \quad 1 \leq k \leq p$$

$$x \in S$$

$$\alpha \in \mathbb{R}$$

onde S é a região admissível definida pelas restrições do problema.

3.2 Projecção dos gradientes das funções objectivo

Uma vez obtidas as restrições activas correspondentes à solução eficiente actual, ou calculada previamente através do programa de realização, podem-se projectar os gradientes das funções objectivo (f.o.) sobre a sub-região não dominada já previamente identificada. Consideram-se activas as restrições cujo índice pertence ao conjunto $I \subset \{1, 2, \dots, m, m+1, \dots, m+n\}$. Para projectar os gradientes das f.o. sobre a sub-região não dominada sob pesquisa, identificada pelas restrições cujo índice pertence ao conjunto $I' \subset I$, obtendo-se direcções tão próximas quanto possível de cada gradiente das f.o., é necessário resolver os seguintes programas lineares:

Para cada $k = 1, \dots, p$

$$\max c^k x$$

sujeito a:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq 0, \quad 1 \leq i \leq m, \quad i \in I'$$

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = 0, \quad 1 \leq i \leq m, \quad i \in I'$$

$$x_j \geq 0, \quad (j+m) \in I'$$

$$x_j = 0, \quad (j+m) \in I'$$

$$x_j \in \mathbb{R}, \quad j = 1, \dots, n.$$

A solução de cada um destes programas lineares é o vector zero, ou o problema é ilimitado. Se todos os problemas resultarem no vector zero, a sub-região I' não pode ser atravessada, isto é não há nenhuma direcção sobre ela que conduza a outra solução eficiente sem passar por soluções não eficientes. Se um programa for ilimitado, a projecção do gradiente da f.o. que lhe corresponde é o vector que conduz à solução infinita e pode ser obtido directamente a partir do quadro do método simplex. Note-se que as restrições de não negatividade são também consideradas como potencialmente activas e que não é necessário alterar o processo acima descrito se a solução eficiente for degenerada. Na prática, às restrições destes programas lineares acrescenta-se uma restrição que limita os problemas e permite utilizar qualquer código do método simplex sem alterações⁷. Trata-se de garantir que o vector x seja normalizado segundo a norma L_1 . Com esta transformação os resultados obtidos são aproximados, garantindo-se no entanto uma boa proximidade de cada projecção ao gradiente da f.o. que o originou.

3.3 Escolha da direcção de pesquisa

Para explorar uma sub-região particular, o agente de decisão tem que escolher uma direcção de pesquisa. Esta direcção pode ser uma combinação convexa do conjunto das projecções dos gradientes das f.o., sobre a sub-região. O processo de escolha da direcção é iterativo, definindo o agente de decisão, em cada iteração, a proximidade mínima aceitável para cada f.o. O processo pode ser descrito da seguinte forma:

- (1) Cálculo do produto interno do gradiente normalizado de cada f.o. com a projecção normalizada respectiva: h_k , $k = 1, \dots, p$. h_k é portanto uma medida da proximidade máxima ao critério k , que se pode obter quando se viaja na sub-região não dominada que se pretende explorar.
- (2) Considerem-se dois conjuntos de índices das funções objectivo: $K = \{1, 2, \dots, p\}$ e $K' = \emptyset$. Ao conjunto K pertencem os índices das f.o. cuja medida de proximidade h ainda não foi restringida pelo agente de decisão e ao conjunto K' pertencem os índices das f.o. em que esta medida já foi restringida.
- (3) Conhecendo h_k , o agente de decisão define g_k , com $k \in K$ e $g_k \leq h_k$. Trata-se da proximidade, mínima aceitável, ao critério k quando se viaja.

⁷ A maior parte dos códigos já existentes para o método simplex não prevê a obtenção de direcções ilimitadas.

(4) Actualização dos conjuntos: $K \leftarrow K \setminus \{k\}$; $K' \leftarrow K' \cup \{k\}$.

(5) Cálculo dos programas lineares seguintes.

Para cada $k \in K$:

$$\max c^k x$$

sujeito a:

$$c^{k'} x \geq g_{k'}, \quad k' \in K'$$

$$x = \sum_{k=1}^p \lambda_k d_k, \quad \lambda_k \geq 0, \quad k = 1, \dots, p.$$

Sendo d_k , $k = 1, \dots, p$, a projecção da k -ésima f.o., obtida segundo a secção 3.2. Para obter as novas direcções à custa das iniciais é necessário que estas últimas sejam geradoras da hiperface em exploração. Com a aproximação prática ao cálculo das direcções iniciais, referida na secção 3.2 (introdução da norma L_1), isto nem sempre se verifica. Pelo que, se se optar por essa restrição, é necessário reprojectar os gradientes das f.o. sobre a hiperface, segundo a secção 3.2.

(6) Recalcular o produto interno do gradiente normalizado de cada f.o., z_k , com $k \in K$, com a sua projecção normalizada, recalculada no ponto (5): h_k , $k \in K$.

(7) Se o agente de decisão pretender modificar a proximidade mínima aceitável da f.o. k' , com $k' \in K'$, então:

$$K' \leftarrow K' \setminus \{k'\}; \quad K \leftarrow K \cup \{k'\}.$$

O agente de decisão pode repetir este passo enquanto $K' \neq \emptyset$, voltando depois ao passo (5).

(8) Se $\#K \neq 1$, volta ao passo (3).

3.4 Viagem sobre a fronteira eficiente

Considere-se que a solução eficiente a deslocar segundo a direcção d (determinada da forma indicada na secção anterior) é x^1 . Então x^2 , definida por $x^2 = x^1 + \Delta d$, é ainda uma solução eficiente se $\Delta \in [\Delta_1, \Delta_2]$, onde:

$$\Delta_1 = \max_j \left(-\frac{x_j^1}{d_j} \right) \text{ se } d_j > 0 \quad \text{e} \quad \Delta_2 = \min_j \left(-\frac{x_j^1}{d_j} \right) \text{ se } d_j < 0$$

Note-se que a direcção tem que ser conhecida também para as variáveis desvio das restrições do problema linear multiobjectivo.

3.5 Determinação do ponto de referência a partir de uma solução não dominada

Seja Cx^1 o ponto não dominado, projecção do ponto de referência q^{*1} segundo a direcção s . Um ponto de referência q^{*2} , segundo a direcção s , cuja projecção é Cx^2 , pode ser o seguinte: $q^{*2} = C(x^2 - x^1) + q^{*1}$.

3.6 Característica dual de uma hiperface eficiente

A partir do quadro do método simplex associado ao programa de realização, que permite projectar um ponto de referência, q^* , sobre a região não dominada (veja-se a secção 3.1), é

possível extrair informação para caracterizar uma hiperface da região admissível. O programa de realização pode ser escrito da seguinte forma:

$$\max \left(\alpha + \frac{\epsilon}{p} \sum_{k=1}^p \frac{c^k x}{s_k} - c^{te} \right)$$

$$c^{te} = \frac{\epsilon}{p} \sum_{k=1}^p \frac{q_k^*}{s_k}$$

sujeito a:

$$c^k x - s_k \alpha - t_k = q_k^*, \quad 1 \leq k \leq p$$

$$x \in S$$

$$\alpha \in \mathbb{R}$$

As restrições $z_k - c^k x = 0$, $1 \leq k \leq p$, são redundantes, isto é, o resultado do programa acima é o mesmo quer elas figurem explicitamente quer não. A introdução explícita das variáveis z , nesta forma, apenas permite obter directamente o valor das funções objectivo.

Seja B uma base associada à solução óptima do programa de realização. Note-se que esta solução óptima pode não corresponder a uma solução básica do programa linear multiobjectivo. Se se expandir a linha dos custos de oportunidade, do quadro do método de simplex, com as linhas associadas às p funções objectivo do programa multiobjectivo obtêm-se os valores das variáveis duais (a menos do sinal, por vezes). $c_B^T B^{-1}$ é a matriz das variáveis duais do problema linear multiobjectivo, transformado no sentido de se obter a solução não dominada que melhor aproxime as aspirações do agente de decisão de acordo com uma determinada escala de valores para cada função objectivo. Não se trata portanto das variáveis duais do problema multiobjectivo inicial. Por exemplo, $W_{ki} = (c_B^T B^{-1})_{ki}$, indica de quanto aumenta a função objectivo k , se se aumentar de uma unidade o termo independente da restrição i (sem sair da mesma base), do problema de realização, isto é, tendo em conta as aspirações e as unidades de escala do agente de decisão. Logo, não se trata aqui do problema dual do problema linear multiobjectivo, mas do problema dual do problema escalarizante de realização.

Considere-se que as variáveis desvio, t_i , $i = 1, \dots, p$, não são básicas, relativamente à base B . Sendo W_{ki} o elemento de $c_B^T B^{-1}$ associado à função objectivo k e à restrição da variável t_i (sem perda de generalidade), ele indicará de quanto aumenta o valor da função objectivo k , quando se aumenta de uma unidade o nível de aspiração da função objectivo i , se não se sair da mesma base. Estes valores das variáveis duais podem servir como característica de uma hiperface da região admissível do problema multiobjectivo, com o intuito de informar o agente de decisão na sua escolha da hiperface não dominada a pesquisar. Para obter esta informação de forma não enviesada é necessário abandonar as unidades de escala, s , das funções objectivo, escolhidas pelo agente de decisão. Repare-se que para viajar sobre a região não dominada, veja-se 3.2, 3.3 e 3.4, as unidades de escala não intervêm e se se pretender caracterizar as diferentes sub-regiões não dominadas com vista a viajar sobre elas, é, portanto, também conveniente abandonar estas unidades.

O acima exposto não depende de B ser ou não uma base admissível (do programa escalarizante de realização de um determinado ponto de referência). Pode-se caracterizar qualquer hiperface eficiente do poliedro admissível do problema original segundo este processo, sem ter que encontrar o ponto de referência cuja projecção resulte numa solução não dominada sobre essa face, isto é, sem ter que encontrar o ponto de referência, que aplicado ao problema de realização torne a base B numa base admissível. Note-se que para uma dada solução não dominada e uma dada unidade de escala, s , existe pelo menos um ponto de referência que lhe corresponde através do programa de realização, se ϵ for suficientemente pequeno (mas positivo). Pode até acontecer trabalhar-se com uma base B , que corresponda a uma solução básica não admissível quando considerado um determinado ponto de referência. No entanto é necessário que a base B cumpra a condição de optimalidade do método simplex, quando aplicado ao programa de realização, garantindo que a hiperface associada é eficiente.

Para caracterizar a hiperface de dimensão $p-1$ a que corresponde a variável desvio r_i (ou várias variáveis desvio⁸, dependendo das dimensões do problema), é necessário encontrar a base B_{r_i} tal que nem as variáveis t_k , $k = 1, \dots, p$, nem a variável r_i (as variáveis desvio correspondentes à hiperface) lhe pertençam. Quando se pretende determinar a característica dual de uma hiperface de dimensão $p-1$ sobre o interior relativo da qual se obteve uma solução não dominada a partir da projecção de um ponto de referência, encontrar esta base torna-se fácil, pois a menos de trocas de variáveis necessárias devido a possíveis degenerescências no programa de realização, a base B possui aquelas características. Noutros casos torna-se necessário encontrar primeiro uma solução não dominada que pertença à hiperface e a partir dela extrair a base com as características referidas.

Se se pretender trabalhar com uma hiperface de dimensão inferior a $p-1$ que contenha no seu interior relativo um ponto não dominado, a sua características dual perde tantos graus de liberdade quanto a diferença que vai da dimensão da hiperface até $p-1$. Isto é, a característica dual indica que se se aumentar de uma unidade um ou outro nível de aspiração (os níveis de aspiração em que isto acontece são determinados pelas características do problema em causa) a proporção relativa de variação dos valores das funções objectivo é a mesma. Por exemplo, se num problema que possuam três funções objectivo ($p = 3$) se tiver uma hiperface eficiente de dimensão um (um segmento de recta), qualquer que seja o nível de aspiração a variar (sendo a variação suficientemente pequena para a projecção se manter sobre essa hiperface) a variação relativa das funções objectivo mantém-se. Isto deve-se ao facto de a base B associada a esta hiperface ter necessariamente que possuir uma ou mais variáveis t 's, para permitir que a face tenha dimensão inferior a $(p-1)$ ⁹.

⁸ r_i é a variável desvio da restrição i do problema original.

⁹ Sendo uma face de dimensão $p' < (p-1)$, existirão $((p-1) - p')$ variáveis t na base B associada à face.

3.7 Determinação das hiperfaces vizinhas da actualmente sob pesquisa

Considere-se que o agente de decisão pesquisa actualmente uma sub-região não dominada correspondente à restrição activa i . Um processo de encontrar as hiperfaces vizinhas da i pode ser o de calcular todos os óptimos alternativos do seguinte problema linear:

$$\begin{aligned} \max & (a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n) \\ \text{s.a: } & x \in S \end{aligned}$$

Encontram-se assim todas as soluções básicas admissíveis da hiperface i . As hiperfaces activas em cada uma destas soluções, serão as vizinhas de i . A partir destas soluções básicas também se extraem as bases necessárias ao cálculo da característica dual, de acordo com a secção 3.6. Se existirem várias (mais do que uma) restrições activas, cada uma delas é tratada de forma independente e depois eliminam-se as bases repetidas que aparecerem.

4. Exemplo Ilustrativo

Nesta secção ilustra-se o atrás exposto recorrendo a um problema muito utilizado na literatura da área, veja-se entre outros Korhonen e Laakso (1986). A fig. 1 representa este problema. O problema linear multiobjectivo é o seguinte:

$$\begin{aligned} \max & z_1 = x_1; \max z_2 = x_2; \max z_3 = x_3 \\ \text{s.a. } & 3x_1 + 2x_2 + 3x_3 \leq 18 \\ & x_1 + 2x_2 + x_3 \leq 10 \\ & 9x_1 + 20x_2 + 7x_3 \leq 96 \\ & 7x_1 + 20x_2 + 9x_3 \leq 96 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

O agente de decisão fixa os seus níveis de aspiração no seguinte ponto de referência: $q^* = (3.5, 4.0, 3.5)$. Eventualmente ele pode requerer alguma informação prévia para escolher o ponto de referência, nomeadamente o ponto ideal ou os pontos óptimos de cada função objectivo individualmente. Neste problema o ponto ideal é o $(6, 4.8, 6)$. O agente de decisão fixa também as unidades de escala para cada função objectivo, $s = (2, 1, 2)$, tendo em atenção se o ponto de referência que escolheu se situa abaixo ou acima do ponto não dominado respectivo. Neste caso o sistema informa que o ponto de referência não é alcançável e portanto as unidades de escala terão um significado inverso¹⁰, isto é, indicam a direcção de degradação do ponto de referência para atingir uma solução não dominada. Assim a primeira e terceira funções objectivo degradar-se-ão o dobro da segunda função objectivo (em princípio). O parâmetro ϵ é fixado em 0.01.

A solução não dominada resultante do problema linear escalarizante de realização é a seguinte:

$$\alpha = -0.833, x_1 = 1.833, x_2 = 3.167, x_3 = 1.833, t_1 = 0, t_2 = 0, t_3 = 0, r_1 = 0.667, r_2 = 0, r_3 = 3.333, r_4 = 3.333.$$

¹⁰ Inverso porque se poderia esperar que quanto maior fosse s_i menor seria a degradação da função objectivo i .

t_k é a variável de desvio da restrição auxiliar associada à k -ésima função objectivo, r_i é a variável de desvio da i -ésima restrição do problema original. Como se pode observar há uma restrição, do problema multiobjectivo, activa: a correspondente à variável r_2 . Determina-se agora a característica dual desta restrição ou sub-região não dominada, veja-se 3.6. A base B , para o problema de realização, considerando as unidades de escala todas iguais à unidade é composta pelas seguintes variáveis: α_2^{11} , x_1 , x_2 , x_3 , r_1 , r_3 , r_4 .

A característica dual, W , da face é obtida a partir das variáveis duais associadas a t_1 , t_2 e t_3 , respectivamente, e de $c_B^T B^{-1}$.

$$W = \begin{bmatrix} 0.75 & -0.5 & -0.25 \\ -0.25 & 0.5 & -0.25 \\ -0.25 & -0.5 & 0.75 \end{bmatrix}$$

Da leitura desta característica dual podem-se tirar, entre outras, as seguintes conclusões:

- Se se aumentar de uma unidade o nível de aspiração da função 1, então as funções objectivo 1, 2 e 3 crescem de 0.75, -0.25 e -0.25, respectivamente, se não se sair da face.
- Trata-se de uma sub-região não dominada interessante se se quiser aumentar o valor das funções objectivo 1 ou 3, pois os custos para as outras duas funções objectivo são relativamente baixos. Se se quiser melhorar o valor da função objectivo 2, nesta sub-região, a degradação das outras duas funções objectivo é relativamente elevada.

Neste ponto o agente de decisão deve pronunciar-se sobre se quer pesquisar com mais detalhe esta sub-região ou se prefere procurar noutras vizinhas desta. Considere-se que o agente de decisão pretende pesquisar esta sub-região detalhadamente.

Primeiro projectam-se os gradientes das funções objectivo sobre a sub-região não dominada, veja-se secção 3.2. Utiliza-se aqui o processo prático, atrás referido, que permite utilizar o código normal do método simplex. Obtêm-se os vectores seguintes, que se podem observar na fig. 1:

$$d_1 = (0.667, -0.333, 0.0); d_2 = (-0.667, 0.333, 0.0) \text{ e } d_3 = (0, -0.333, 0.667)$$

Apresenta-se o resultado do produto interno de cada um destes vectores d , com o gradiente da função objectivo respectiva, isto é os h_k . Veja-se tabela I, interacção A.

	h_1	h_2	h_3
A	0.667	0.333	0.667
B	0.667	-0.333	0.000
B1	0.300	0.133	0.367
C	0.300	-0.333	0.300

Tabela I - Proximidades máximas possíveis durante o deslocamento

¹¹ Repare-se que se fez a mudança de variável $\alpha = \alpha_1 - \alpha_2$, $\alpha_1 \geq 0$ e $\alpha_2 \geq 0$, devido a α ser uma variável sem restrição de sinal.

O agente de decisão fixa como proximidade mínima admissível entre a função objectivo um e a direcção de pesquisa o valor 0.666, $g_1 \geq 0.666$. Volta-se a projectar os gradientes das funções objectivo, mas apenas para a segunda e terceira funções. A cada um dos programas lineares necessários acrescenta-se a restrição $z_1 \geq 0.666$, isto é $x_1 \geq 0.666$. Como resultado obtêm-se os vectores seguintes:

$$d_2 = (0.667, -0.333, 0.0) \text{ e } d_3 = (0.667, -0.333, 0.0).$$

O resultado do produto interno de cada um destes vectores d com o gradiente da função objectivo respectiva encontra-se na tabela I, interacção B. Repare-se que para a função objectivo 1, o resultado está a escuro pois esta proximidade já foi fixada.

O agente de decisão não está satisfeito pois também pretendia aumentar o valor da função objectivo 3! Face aos resultados da tabela I, ele vai voltar atrás, relaxando a proximidade mínima aceitável para a função objectivo 1. Passa a ser de $g_1 \geq 0.300$. Projectam-se de novo os gradientes das funções objectivo e obtêm-se:

$$d_2 = (0.300, 0.133, -0.567) \text{ e } d_3 = (0.300, -0.333, 0.367).$$

Na tabela I, iteração B1 surge a proximidade máxima resultante. O agente de decisão fixa agora $g_3 \geq 0.300$ e obtêm-se o vector seguinte:

$$d_2 = (0.333, -0.333, 0.333).$$

É este o vector de pesquisa, d, face às preferências locais do agente de decisão. Para se poder viajar na região não dominada é necessário encontrar o valor das variáveis desvio da direcção. A direcção de pesquisa (representada na fig. 1) é:

$$d = (0.333, -0.333, 0.333, -1.333, 0, 1.333, 1.333).$$

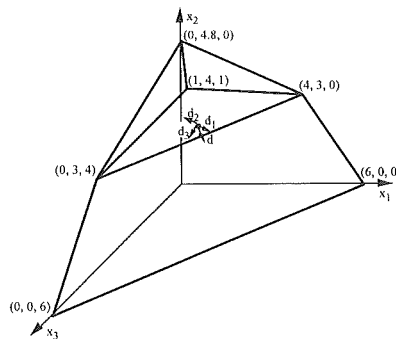


Fig. 1 - Direcção de pesquisa sobre a região não dominada

Torna-se agora necessário saber de quanto se pode viajar sobre a sub-região eficiente, segundo esta direcção. De acordo com a secção 3.4 calcula-se:

$$\Delta_1 = \max\left(\frac{1.833}{-0.333}, \frac{1.833}{0.333}, \frac{3.333}{-1.333}, \frac{3.333}{1.333}\right) = -2.5 \quad \Delta_2 = \min\left(\frac{3.166}{-0.333}, \frac{0.666}{-1.333}\right) = +0.5$$

$$A \in [-2.5, 0.5]$$

Sendo o ponto inicial, p, igual a

$$x_1 = 1.833, x_2 = 3.167, x_3 = 1.833, r_1 = 0.667, r_2 = 0, r_3 = 3.333, r_4 = 3.333,$$

obtêm-se um ponto final no sentido inverso, p^1 , igual a

$$x_1 = 1, x_2 = 4, x_3 = 1, r_1 = 4, r_2 = 0, r_3 = 0, r_4 = 0,$$

e um ponto final no sentido directo, p^2 , igual a

$$x_1 = 2, x_2 = 3, x_3 = 2, r_1 = 0, r_2 = 0, r_3 = 3.5, r_4 = 3.5.$$

A fig. 2 apresenta o gráfico de variação do valor das funções objectivo ao longo da direcção de pesquisa. O agente de decisão desejava melhorar o valor das funções objectivo 1 e 3, por isso o sentido que lhe interessa é o que conduz a p^2 . Repare-se que de acordo com a secção 3.5, o ponto de referência que conduziria a esta solução não dominada, para as mesmas unidades de escala, seria: $p^{2*} = (3.667, 3.833, 3.667)$.

Pretende-se, agora, verificar se vale a pena continuar a pesquisa na nova face a que se chegou. Para isso é necessário obter a característica dual desta nova face. A partir de p^2 , verifica-se que a base B_{r_1} , com que se vai trabalhar é não dominada e constituída por: $\alpha_2, x_1, x_2, x_3, r_2, r_3, r_4$. Aplicando o que se descreve na secção 3.6, obtém-se a seguinte matriz:

$$W = \begin{bmatrix} 0.625 & -0.25 & -0.375 \\ -0.375 & 0.75 & -0.375 \\ -0.375 & -0.25 & 0.625 \end{bmatrix}$$

Repare-se que parece uma sub-região não dominada interessante para aumentar a função objectivo dois. No entanto, como esta função vem a degradar-se ao longo do trajecto efectuado, nesta face a tendência natural será a de piorar essa função.

O agente de decisão pretende atravessar a aresta! Existem agora duas restrições activas: a associada a r_2 (já estava activa antes de atingir a aresta) e a associada a r_1 (sub-região a pesquisar). Inicia-se outra vez o processo de procura da direcção de deslocamento, projectando-se os gradientes das funções objectivo sobre a nova sub-região não dominada. Obtém-se os seguintes vectores:

$d_1 = (0.5, 0, -0.5)$ Nesta direcção a variável desvio associada à segunda restrição é nula (o cálculo não é apresentado), o que significa que esta direcção está sobre a aresta.

$d_2 = (0, 0, 0)$ Isto significa que não se pode melhorar o valor da segunda função objectivo nesta face, a partir da solução não dominada p^2 .

$d_3 = (-0.5, 0, 0.5)$ Mais uma vez a direcção está sobre a aresta.

Note-se que a combinação linear de d_1 e d_3 não permite explorar a face associada a r_1 , pois ambas as direcções estão sobre a mesma aresta. Isto deve-se à modificação do programa de projecção dos gradientes a que se refere a secção 3.2: para que as soluções não fossem infinitas obrigou-se a que a norma L_1 da direcção seja igual à unidade. Uma vez que os gradientes são reprojectados em cada interacção isto não impede a exploração da face, aliás como se irá ver na continuação do exemplo.

Ao agente de decisão apresenta-se o resultado do produto interno de cada um destes vectores d , com o gradiente da função objectivo respectiva. Veja-se na tabela II, interacção A.

Fixa-se como proximidade mínima admissível entre a função objectivo um e a direcção de pesquisa o valor 0.25, $g_1 \geq 0.25$. Obtém-se os seguintes vectores:

$$d_2 = (0.5, 0, -0.5) \text{ e } d_3 = (0.25, -0.6, 0.15).$$

	h_1	h_2	h_3
A	0.5	0.0	0.5
B	0.25	0.0	0.15
C	0.25	-0.6	0.15

Tabela II - Proximidades máximas possíveis durante o deslocamento

O agente de decisão fixa como proximidade mínima admissível entre a função objectivo três e a direcção de pesquisa o valor 0.15, $g_3 \geq 0.15$, obtendo-se: $d_2 = (0.25, -0.6, -0.15)$. É esta a direcção de pesquisa. Quando aumentada com as variáveis desvio aparece como: $d = (0.25, -0.6, 0.15, 0, 0.8, 8.7, 8.9)$. Face a esta direcção $\Delta_1 = 0$ e $\Delta_2 = 5$.

O agente de decisão pode, portanto, deslocar-se desde o ponto p^2 até ao ponto p^3 , cujas coordenadas são:

$$x_1 = 3.25, x_2 = 0, x_3 = 2.75, r_1 = 0, r_2 = 4, r_3 = 47, r_4 = 48.$$

Repare-se que com esta direcção e o ponto p^2 não há lugar a deslocações no sentido inverso.

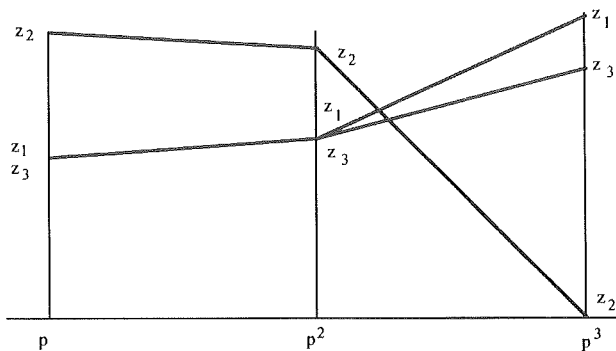


Fig. 2 - Trajecto ao longo das direcções de pesquisa

O processo aqui exemplificado podia repetir-se até o agente de decisão se satisfazer com uma determinada solução.

5. Conclusão

Neste artigo apresentou e exemplificou-se um método de programação linear multiobjectivo que se integra na grande família dos métodos baseados em pontos de referência, tal como a secção 2 refere. Com este novo método procurou-se essencialmente melhorar as ferramentas de controlo ao dispor do agente de decisão quando realiza pesquisas sobre a região não dominada de um problema. Estas pesquisas podem ter o intuito de seleccionar a solução preferida do agente de decisão ou apenas o de conhecer melhor a região referida. As ferramentas que o método põe à disposição permitem, por um lado, exercer um controlo elevado sobre a direcção de pesquisa local de uma solução não dominada a partir de outra e, por outro lado, apreender facilmente as características das diferentes sub-regiões não dominadas, no que diz respeito às possibilidades de variação linear relativa. É, portanto, um método que, a par de apresentar uma evolução considerável, face aos métodos referidos na secção 2, no que diz respeito ao controlo da direcção de pesquisa, apresenta ainda possibilidades de apoiar um tipo de pesquisa não local,

sem contudo ser puramente estratégica. Isto é, através da característica dual das diferentes hiperfaces não dominadas, surgindo à medida que se percorre a região não dominada, consegue-se ultrapassar algumas delas com perspectivas pouco interessantes para as preferências do agente de decisão, acelerando-se portanto a pesquisa local com alguma informação estratégica.

6. Referências

- [1] Benayoun, R., Montgolfier, J., Tergny, J. e Larichev, O., *Linear programming with multiple objectives: step method (STEM)*, Mathematical Programming 1 (1971) 366-375.
- [2] Costa, J.P. e Clímaco, J., *A multiple reference point parallel approach in MCDM*, in 'Multiple Criteria Decision Making: Expand and Enrich the Domain of Thinking and Application', G.H. Tzeng, H.F. Wang, U.P. Wen e P.L. Yu (Eds), Springer Verlag (1994) 255-263.
- [3] Halme, M., *Local characterization of efficient solutions in interactive multiple objective linear programming*, Tese de doutoramento, Acta Academiae Oeconomical Helsingiensis, Series A:84, The Helsinki School of Economics and Business Administration (1992).
- [4] Halme, M. e Korhonen, P., *Nondominated tradeoffs and termination in multiple objective linear programming*, in 'Improving decision making in organizations', A.G. Lockett e G. Islei (Eds), Springer-Verlag, LNEMS 385 (1989) 410-423.
- [5] Iserman, *Proper efficiency and the linear vector maximum problem*, Operations Research 22 (1974) 189-191.
- [6] Korhonen, *VIG - A visual interactive approach to goal programming. User's guide. Version 2.20-99*, NumPlan (1987).
- [7] Korhonen, P. e Laakso, J., *A visual interactive method for solving the multiple criteria problem*, in 'Interactive decision analysis', M. Grauer e A.P. Wierzbicki (Eds), Springer-Verlag, LNEMS 229 (1984) 146-153.
- [8] Korhonen, P. e Laakso, J., *A visual interactive method for solving the multiple criteria problem*, European Journal of Operational Research 24 (1986) 277-287.
- [9] Korhonen, P. e Wallenius, J., *A Pareto Race*, Naval Research Logistics 35 (1988) 615-623.
- [10] Korhonen, P. e Wallenius, J., *A computer graphics-based decision support system for multiple objective linear programming*, European Journal of Operational Research 60 (1992) 280-286.
- [11] Lewandowski, A. e Wiersbicki, A., *Aspiration based decision analysis and support. Part I: theoretical and methodological backgrounds*, IIASA wp-88-03 (1988).
- [12] Lewandowski, A., Kreglewski, T., Rogowski, T. e Wiersbicki, A., *Decision support systems of DIDAS family (Dynamic Interactive Decision Analysis and Support)*, em 'Theory, software and testing examples in decision support systems', A. Lewandowski e A. Wierz (Eds), IIASA wp-88-071 (1988) 18-41.
- [13] Nakayama, H. e Sawaragi, *Satisficing trade-off method for multiobjective programming*, em 'Interactive decision analysis', M. Grauer e A. Wierzbicki (Eds), Springer-Verlag, LNEMS 229 (1984) 113-122.
- [14] Steuer, R., *Multiple criteria optimization: theory, computation and application*, John Wiley & Sons (1986).
- [15] Steuer, R. e Choo, E.-U., *An interactive weighted Tchebycheff procedure for multiple objective programming*, Mathematical Programming 26 (1983) 326-344.
- [16] Vincke, Ph., *Une méthode interactive en programmation linéaire à plusieurs fonctions économiques*, Revue Française d'Informatique et de Recherche Opérationelle 2 (1976) 5-20.

A IMPORTÂNCIA DA PESQUISA UNIDIMENSIONAL NA EFICIÊNCIA DOS MÉTODOS DE OPTIMIZAÇÃO

João L.H. Matias

Departamento de Matemática
Universidade de Trás-os-Montes e Alto Douro
5000 Vila Real - Portugal

Edite M.G.P. Fernandes

Departamento de Produção e Sistemas
Universidade do Minho
4710 Braga - Portugal

Abstract

In nonlinear multidimensional optimization line search techniques aim to compute an optimum value for the step length to take along the search direction. The use of line search is related to the globalization of the algorithm.

In this paper eight line search techniques are tested. Six of them are defined by a search step followed by a local approximation of the function. Davies, Swann and Campey search technique is implemented in four and Golden Section in two of them. Quadratic and cubic approximations are then introduced. Finally Armijo and Curry-Altman-Armijo criteria complete the set. Goldstein-Armijo condition is then used to guarantee a sufficient decrease of the function.

Numerical results show that Armijo and Curry-Altman-Armijo criteria are more efficient in the sense that they need less iterations and functions calculations. Between approximation techniques, cubic based interpolation is the most efficient.

Resumo

As técnicas de pesquisa unidimensional, em optimização multidimensional e não linear, têm como objectivo calcular um valor óptimo para o comprimento do passo a realizar ao longo de uma direcção de procura. A importância da introdução da técnica de procura unidimensional está relacionada com a globalização do algoritmo de optimização não linear.

Os oito algoritmos de pesquisa unidimensional testados neste artigo, são de dois tipos. Seis deles pertencem à classe dos métodos mistos de procura e aproximação local. Destes, quatro determinam um intervalo que contém o mínimo, baseado no esquema de procura de Davies, Swann e Campey e dois deles no esquema da Secção Dourada. A aproximação local que é feita no intervalo é baseada ou na interpolação quadrática ou na cúbica. Os dois últimos definem uma sucessão de valores que satisfazem propriedades muito simples e são conhecidos por critérios de Armijo e Curry-Altman-Armijo. A condição de Goldstein-Armijo é usada para garantir um decréscimo significativo do valor da função.

Os resultados mostram que os critérios de Armijo e Curry-Altman-Armijo são mais eficientes, no sentido de que necessitam de menos iterações e de menos cálculos da função objectivo do que os outros. Dos algoritmos que incorporam uma fase de aproximação, os baseados na aproximação cúbica são mais eficientes.

Keywords

Unconstrained nonlinear optimization, line search.

1. Introdução

Dos problemas de optimização não lineares, quicá os mais complexos pela natureza das funções envolvidas, trataremos apenas dos problemas sem restrições nas variáveis que são formulados da seguinte maneira

$$\text{minimizar}_{x \in \mathbb{R}^n} f(x), \quad (1)$$

em que $f(x)$ é uma função real e não linear em x . Esta formulação também contempla problemas com qualquer tipos de restrições, uma vez que as técnicas de penalização e os métodos de pontos interiores transformam um problema com restrições, noutro, sem restrições nas variáveis.

Para a resolução do problema (1), os métodos do gradiente de primeira ordem, por necessitarem apenas de informação relativa ao vector gradiente das primeiras derivadas da função, são muitas vezes usados em detrimento dos de segunda ordem que exigem cálculos de segundas derivadas e resolução de sistemas lineares. Como exemplos dos de segunda ordem são conhecidos o método de Newton ([7] e [12]), a modificação de Murray baseada na decomposição Cholesky ([21]), o algoritmo de segurança de Newton ([6] e [21]), o método de Newton modificado ([4] e [15]), a aproximação de Newton baseada em diferenças finitas ([12]) e o método de Greenstadt ([21]) que necessita do cálculo dos valores e vectores próprios.

Quer os métodos de primeira ordem quer os de segunda são iterativos e geram uma sucessão de aproximações $\{x^{(k)}\}$ à solução, a partir de um valor inicial, $x^{(0)}$. Se o processo for convergente, tem-se

$$\lim_{k \rightarrow \infty} x^{(k)} = x^*,$$

sendo x^* a solução de (1).

Na sua essência os métodos iterativos calculam direcções de procura, $d^{(k)}$, que quando adicionadas à aproximação $x^{(k)}$, da iteração k , geram outra aproximação $x^{(k+1)}$,

$$x^{(k+1)} = x^{(k)} + d^{(k)}, \quad (2)$$

para a iteração seguinte.

Além da direcção de procura, surge ainda outra questão primordial à qual há que dar resposta:

- De acordo com algum critério de medida, qual é o comprimento do passo, $\alpha^{(k)}$, a efectuar ao longo da referida direcção,

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)}d^{(k)}, \quad (3)$$

por forma que a próxima aproximação possa ser considerada melhor do que a anterior?

O cálculo desse comprimento é conhecido na literatura por pesquisa unidimensional.

Assim, para uma descrição mais detalhada das questões até agora levantadas, determinação de uma direcção de procura e cálculo do comprimento do passo a efectuar ao longo da direcção, o artigo foi estruturado da seguinte maneira: Na Secção 2 faz-se uma descrição de um método do gradiente de primeira ordem, conhecido por método Quasi-Newton. A Secção 3 foi reservada para as técnicas de pesquisa unidimensional e são descritos oito algoritmos. As

primeiras subsecções descrevem as técnicas de procura de Davies, Swann e Campey e da Secção Dourada bem como dois tipos de interpolação polinomial, designadamente a quadrática e a cúbica. Os dois algoritmos baseados no critério de Armijo surgem no final da secção. Finalmente, os resultados dos testes numéricos realizados, com o objectivo de determinar o desempenho das técnicas descritas na secção anterior, são apresentados na Secção 4. As conclusões deste estudo comparativo surgem na última secção, bem como sugestões, no âmbito de possíveis melhoramentos a introduzir nos algoritmos, são apresentadas no final do artigo.

2. Direcções de Procura

Vamos designar o vector gradiente de f por $g(x)$, a matriz Hessiana de $f(x)$ por $G(x)$, e para o mínimo x^* , verificam-se as seguintes condições suficientes:

$$\|G(x^*)\| = 0 \quad \text{e} \quad p^T G(x^*) p > 0 \tag{4}$$

para qualquer vector de $p \in \mathbb{R}^n \setminus \{0\}$. A primeira condição é de primeira ordem e define pontos estacionários. Nesta secção descreve-se um método que apenas requer informação relativa às primeiras derivadas da função f e possui convergência local superlinear.

2.1 Equações de Newton

Embora não seja o método testado neste trabalho, optou-se por fazer uma breve introdução ao método de Newton, pela sua relevância e ligação com o método Quasi-Newton.

O desenvolvimento em série de Taylor da função objectivo, até às derivadas de segunda ordem, para aproximar o valor mínimo de f , a partir de um ponto x , próximo de x^* , origina

$$f(x+d) \approx f(x) + g(x)^T d + \frac{1}{2} d^T G(x) d$$

em que o lado direito da equação define um modelo quadrático que aproxima a função no ponto x .

Derivando esta expressão em ordem aos elementos d_j de d , considerando g e G fixos, tem-se

$$\frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^n d_j \frac{\partial^2 f(x)}{\partial x_j \partial x_i} = 0, \quad i = 1, 2, \dots, n$$

ou na forma matricial

$$G(x)d = -g(x). \tag{5}$$

Como a função objectivo f não é quadrática, o novo ponto obtido a partir de x , $x+d$, sendo d calculado de (5), não é necessariamente o mínimo desejado, havendo necessidade de repetir o processo iterativamente, ou seja,

$$x^{(k+1)} = x^{(k)} + d^{(k)} \quad \text{com} \tag{6}$$

$$G(x^{(k)})d^{(k)} = -g(x^{(k)}) \quad \text{para} \quad k = 0, 1, 2, \dots \tag{7}$$

em que $x^{(0)}$ é uma aproximação inicial. A equação (7) é conhecida por equação Newton.

Se a aproximação inicial utilizada estiver suficientemente próxima da solução, a sucessão de pontos gerados pelo presente método convergirá para x^* quadraticamente. Fora da região de

convergência, o método de Newton pode falhar, podendo mesmo obter-se matrizes Hessianas singulares.

2.2 Método Quasi-Newton

O problema da singularidade da matriz do sistema (7) e o aparecimento de matrizes indefinidas, ou semi-definidas negativas, podem ser fácil e economicamente ultrapassados com a utilização de uma aproximação à matriz Hessiana que seja simétrica e definida positiva. Quando a função $f(x)$ tem uma forma complicada, o cálculo das segundas derivadas pode ser demasiado complexo e laborioso. Além disso, quando o problema tem uma dimensão razoável, a matriz Hessiana exige o cálculo de $\frac{n(n+1)}{2}$ segundas derivadas diferentes, e a implementação torna-se muito dispendiosa. Os métodos do tipo Quasi-Newton surgiram como uma solução aos problemas referidos.

Na descrição que se segue as aproximações á matriz G e à sua inversa, são representadas, respectivamente, por B e H . O vector direcção, em cada iteração, é calculado usando a matriz B no sistema (7),

$$B^{(k)}d^{(k)} = -g(x^{(k)}) \quad k = 0, 1, \dots$$

ou, de modo equivalente, usando H na equação

$$d^{(k)} = -H^{(k)}g(x^{(k)}) \quad k = 0, 1, \dots \quad (8)$$

Optando a partir de agora pela formulação (8) é necessário conhecer uma estimativa inicial da matriz de aproximação, $H^{(0)}$, bem como o processo que actualiza a matriz $H^{(k)}$ de iteração para iteração.

Admitindo que a função objectivo possui derivadas de segunda ordem contínuas, e usando o desenvolvimento em série de Taylor do gradiente de f , obtém-se,

$$g(x^{(k+1)}) \approx g(x^{(k)}) + G(x^{(k)})(x^{(k+1)} - x^{(k)})$$

ou seja,

$$G(x^{(k)})s^{(k)} \approx y^{(k)} \quad (9)$$

em que $s^{(k)} = \alpha^{(k)}d^{(k)} = x^{(k+1)} - x^{(k)}$ e $y^{(k)} = g(x^{(k+1)}) - g(x^{(k)})$ define a variação calculada no vector gradiente.

Sendo desejável que a matriz $H^{(k)}$ estime o melhor possível $G^{-1}(x^{(k)})$, deve exigir-se, de (9), que

$$s^{(k)} = H^{(k)}y^{(k)}$$

No entanto, para verificar esta condição é necessário conhecer $x^{(k+1)}$ e $g(x^{(k+1)})$ e estes valores só serão calculados depois de resolver (8) já com $H^{(k)}$. Assim, uma das condições que devem satisfazer as fórmulas de actualização de $H^{(k)}$ é a determinada pela equação

$$s^{(k)} = H^{(k+1)}y^{(k)} \quad (10)$$

e que é conhecida por condição Quasi-Newton.

Se o método for convergente, é natural esperar-se que $H^{(k)}$ convirja para $G(x^*)^{-1}$ quando $k \rightarrow \infty$. Embora a matriz de actualização $E^{(k)}$, a partir da qual se obtém $H^{(k+1)} = H^{(k)} + E^{(k)}$,

não seja única, pode-se provar ([7], [9] e [21]) que todas as fórmulas de actualização de $H^{(k)}$, que satisfazem a condição Quasi-Newton, são da forma

$$H^{(k+1)} = H^{(k)} + s^{(k)}u^{(k)T} - H^{(k)}y^{(k)}v^{(k)T} \tag{11}$$

em que $u^{(k)}$ e $v^{(k)}$ são vectores não nulos, tais que

$$u^{(k)T}y^{(k)} = 1$$

e

$$v^{(k)T}y^{(k)} = 1.$$

De facto

$$H^{(k+1)}y^{(k)} = H^{(k)}y^{(k)} + s^{(k)}u^{(k)T}y^{(k)} - H^{(k)}y^{(k)}v^{(k)T}y^{(k)} = s^{(k)},$$

ou seja, satisfazem a condição (10).

Se

$$u^{(k)} = \frac{s^{(k)}}{s^{(k)T}y^{(k)}}$$

e

$$v^{(k)} = \frac{H^{(k)}y^{(k)}}{y^{(k)T}H^{(k)}y^{(k)}}$$

obtém-se de (11) a fórmula de actualização

$$H^{(k+1)} = H^{(k)} + \frac{s^{(k)}s^{(k)T}}{s^{(k)T}y^{(k)}} - \frac{(H^{(k)}y^{(k)})(H^{(k)}y^{(k)})^T}{y^{(k)T}H^{(k)}y^{(k)}} \tag{12}$$

que é conhecida por fórmula DFP, devida a Davidon, Fletcher e Powell, em que

$$s^{(k)} = -\alpha^{(k)}H^{(k)}g(x^{(k)}), \quad k = 0, 1, 2, \dots$$

Se a matriz $H^{(k)}$ for definida positiva, qualquer que seja k , a direcção de procura não é ortogonal ao gradiente, e os denominadores da expressão (12) nunca se anulam. Na prática, o facto da função não ser unimodal ao longo da direcção de procura, a existência de pontos de descanso¹ e o aparecimento de erros de arredondamento, podem originar uma matriz $H^{(k)}$ não definida positiva, causando a divergência do algoritmo. Uma forma de ultrapassar alguns destes problemas consiste em reinicializar a matriz H com a identidade, tal como se faz no início do processo. Se o número de variáveis for n , a reinicialização de H , pode ser feita de n em n iterações.

Existem outras fórmulas de actualização da matriz $H^{(k)}$. Sugere-se a consulta dos trabalhos de Abdy e Dempster em [1], Fernandes em [6], Fletcher em [7] e Gill, Murray e Wright em [9].

Resultados referentes às propriedades locais deste tipo de métodos asseguram a existência de uma vizinhança da solução e de valores de certos parâmetros que garantem a sua convergência para a solução. Relativamente à razão de convergência sabe-se já, hoje em dia, que este tipo de métodos tem convergência q-superlinear. Os estudos teóricos realizados neste sentido (por exemplo [11], [12] e [16]) apontam para a necessidade da direcção $d^{(k)} = -H^{(k)}g(x^{(k)})$ aproximar a direcção Newton assintoticamente. A condição de

¹ Um ponto de descanso é um ponto estacionário que não é máximo nem mínimo.

consistência, que exige que a sucessão $\{H^{(k)}\}$ convirja para $G^{-1}(x^*)$, não é necessária para a prova de convergência.

3. Pesquisa Unidimensional

Nesta secção faz-se uma descrição dos algoritmos usados na pesquisa unidimensional. Entende-se por pesquisa unidimensional, o processo que pesquisa, ao longo de uma direcção de procura, o valor óptimo do escalar α que nos dá o comprimento do passo a executar (ver [17]). A direcção de procura pode ser calculada pelo método referido na secção anterior.

A importância da pesquisa unidimensional está relacionada com o facto de ser vantajoso garantir convergência do método para a solução x^* , qualquer que seja a aproximação inicial atribuída ao método iterativo. Chama-se a isto, a globalização do método. Os estudos referentes à convergência q-superlinear do método Quasi-Newton, bem como à convergência q-quadrática do método de Newton, determinam a existência de uma vizinha da solução, onde é possível, e só aí, garantir a referida convergência. Assim, estes métodos são identificados como localmente convergentes. A globalização define o passo seguinte no estudo teórico de um método e consolida a eficiência e a robustez da sua implementação.

A globalização pode ser feita recorrendo-se a duas técnicas diferentes. A primeira, conhecida por procura unidimensional, foi a seleccionada para este trabalho, pela sua simplicidade. A outra, conhecida por técnica das regiões de confiança é hoje em dia muito utilizada e tem uma ligação estreita com os problemas com restrições nas variáveis ([2]). O trabalho de Vicente [20] apresenta uma comparação entre estas técnicas a nível das condições que garantem a globalização do método.

Das técnicas de pesquisa unidimensional existem duas abordagens. A primeira, que adoptámos neste artigo força um comportamento monótono decrescente da função, ao longo do processo iterativo, tentando verificar a condição (13) ou as condições (28) e (29) (na Secção 4). A segunda não exige um decréscimo do valor da função em todas as iterações. A técnica é conhecida por estratégia não monótona de estabilização. Recomendamos os trabalhos de Grippo, Lampariello e Lucidi [10] e Soares [19] para uma descrição e consulta de bibliografia sobre o assunto.

Em todos os algoritmos de pesquisa unidimensional apresentados neste capítulo, admite-se que o vector direcção, $d^{(k)}$, e a aproximação $x^{(k)}$, de x^* , são conhecidos na iteração k . Assim, o objectivo é calcular o valor $\alpha^{(k)}$ que minimiza localmente $f(x^{(k)} + \alpha d^{(k)})$, isto é,

$$\alpha^{(k)} = \arg \min_{\alpha} f(x^{(k)} + \alpha d^{(k)}). \quad (13)$$

Pelo facto de f ser uma função não linear em x , esta condição equaciona um problema de minimização não linear da função f , a uma dimensão, que só por si é um problema complicado. A solução analítica só em casos particulares de funções polinomiais de grau quatro, três ou dois pode ser calculada. Na generalidade, só é possível calcular aproximações à solução do problema unidimensional. Os métodos de aproximação, baseados em formas polinomiais de

grau baixo, designadamente de grau dois ou três são os mais usados. O mínimo da forma polinomial será usado como aproximação à solução pretendida.

Este raciocínio é a ideia base dos métodos de aproximação que focaremos nesta Secção em 3.1.1, 3.1.2 e 3.3.1.

Uma vez que a condição (13) só raramente é verificada, surge a necessidade de utilizar outras condições que garantam que o valor calculado $\alpha^{(k)}$, origina uma nova aproximação $x^{(k+1)}$, melhor que a anterior, $x^{(k)}$. Melhor significa, neste problema, que o valor da função diminui, de uma certa quantidade, de iteração para iteração. Em 3.4 e 3.5 apresentam-se dois critérios baseados nessas condições.

Os métodos de aproximação são normalmente bem sucedidos se o intervalo, definido pelos valores usados na aproximação de f (três no caso da aproximação quadrática e quatro ou dois, no caso da cúbica), contiver o mínimo pretendido. Este problema do encaixe do mínimo num intervalo de valores é importante e pode ser facilmente resolvido se for utilizado um dos três métodos descritos em 3.1, 3.2 e 3.3. O primeiro é o método de Davies, Swann e Campey e o segundo é conhecido por método da Secção Dourada. O método de Davies, Swann e Campey modificado, em 3.3, definirá também um intervalo que contém o mínimo para a interpolação cúbica baseada apenas em dois pontos, mas utiliza informação relativa a f' , para esse fim.

Começemos pelo método de Davies, Swann e Campey.

3.1 Algoritmo de Davies, Swann e Campey

Este algoritmo é caracterizado na literatura por ser misto ([1]). Tem uma fase de procura do intervalo que contém o mínimo e a outra, de aproximação, que pode ser baseada em interpolação quadrática ou cúbica.

Partindo de uma aproximação inicial, α_0 , a estratégia do algoritmo de Davies, Swann e Campey (DSC), consiste em definir, um conjunto de pontos, separados uns dos outros sucessivamente de δ , 2δ , 4δ , 8δ , ..., por forma a conseguir um intervalo que encaixe o mínimo. Isto acontece quando se encontra, nesse conjunto, o primeiro ponto para o qual a função sofre um aumento, em relação ao ponto anterior.

Os incrementos (múltiplos de δ) são sucessivamente adicionados ou subtraídos ao ponto inicial, originando respectivamente uma procura no sentido positivo ou negativo, desde que a função esteja a diminuir. O processo pára quando é encontrado um ponto, para o qual a função tenha aumentado de valor, quando comparado com o valor da função no ponto anterior, isto é,

$$\begin{array}{ll} \alpha_1 = \alpha_0 + \delta & \alpha_1 = \alpha_0 - \delta \\ \alpha_2 = \alpha_1 + 2\delta & \alpha_2 = \alpha_1 - 2\delta \\ \alpha_3 = \alpha_2 + 4\delta & \text{ou} \quad \alpha_3 = \alpha_2 - 4\delta \\ \alpha_4 = \alpha_3 + 8\delta & \alpha_4 = \alpha_3 - 8\delta \end{array}$$

...

...

até $f(\alpha_k) > f(\alpha_{k-1})$. Em ambos os casos, os três últimos pontos, definem um intervalo que contém o mínimo. Assim, tem-se $\alpha_{k-2} < \alpha_{k-1} < \alpha_k$ se a procura foi no sentido positivo ou

$\alpha_k < \alpha_{k-1} < \alpha_{k-2}$ se foi no sentido negativo, com $f(\alpha_{k-2}) \geq f(\alpha_{k-1})$ e $f(\alpha_{k-1}) < f(\alpha_k)$. O último subintervalo, tem uma amplitude dupla da amplitude do subintervalo anterior. Calculado o ponto médio deste último subintervalo, obtêm-se quatro pontos igualmente distanciados, sendo a distância entre eles a amplitude do penúltimo subintervalo calculado.

3.1.1 Aproximação quadrática

Quando se conhece um intervalo que contém o mínimo, a maneira mais fácil de calcular uma aproximação a esse mínimo consiste em aproximar a função dada, $f(\alpha)$, por uma forma polinomial de grau baixo. O mínimo da forma polinomial é então usado como aproximação ao mínimo procurado.

Para construir um polinómio interpolador do segundo grau precisamos de três pontos. Dos quatro pontos igualmente distanciados obtidos pela fase de procura do DSC, escolhemos três. Por análise do valor da função nos dois pontos interiores obtidos pela procura DSC, escolhe-se o que tem menor valor da função. Este e os adjacentes definem o terno.

Seja

$$p(\alpha) = a\alpha^2 + b\alpha + c \quad a, b, c \in \mathbb{R}, a \neq 0$$

o polinómio do segundo grau, em que a , b e c , são constantes a determinar.

O mínimo do polinómio quadrático, verifica as condições

$$p'(\alpha) = 2a\alpha + b = 0,$$

cuja solução é

$$\hat{\alpha} = \frac{-b}{2a}, \quad (14)$$

e

$$p''(\alpha) = 2a > 0.$$

Considerando $\alpha_1 < \alpha_2 < \alpha_3$, o cálculo das constantes reais a , b e c , pelo método dos coeficientes indeterminados envolve a resolução de um sistema linear de três equações, obtendo-se por substituição da sua solução (14)

$$\hat{\alpha} = \frac{1}{2} \frac{(\alpha_2^2 - \alpha_3^2)f_1 + (\alpha_3^2 - \alpha_1^2)f_2 + (\alpha_1^2 - \alpha_2^2)f_3}{(\alpha_2 - \alpha_3)f_1 + (\alpha_3 - \alpha_1)f_2 + (\alpha_1 - \alpha_2)f_3} \quad (15)$$

Como os pontos (α_1, f_1) , (α_2, f_2) e (α_3, f_3) verificam

$$2(\alpha_2 - \alpha_3) = \alpha_3 - \alpha_1 = 2(\alpha_1 - \alpha_2) = 2h$$

a fórmula (15) pode ser simplificada, obtendo-se

$$\hat{\alpha} = \alpha_2 + \frac{h}{2} \frac{f_1 - f_3}{f_1 - 2f_2 + f_3}. \quad (16)$$

Este mínimo do polinómio quadrático não é necessariamente o mínimo da função $f(\alpha)$. O processo terá de ser repetido, desde a fase de procura, considerando agora como aproximação inicial, $\alpha_0 = \hat{\alpha}$, esta solução encontrada. O valor do incremento δ , deve também sofrer uma diminuição, isto é, $\delta = \mu\delta$, com $\mu < 1$. Detalhes da implementação do correspondente algoritmo podem ser encontrado em [1] e [14].

3.1.2 Aproximação cúbica

Para construir um polinómio cúbico são necessário quatro pontos, se apenas se conhecerem os valores que a função toma nesses pontos. Terminada a fase de procura do algoritmo de Davies, Swann e Campey ficamos com um intervalo que contém o mínimo.

Seja

$$p(\alpha) = a\alpha^3 + b\alpha^2 + c\alpha + d \quad a, b, c, d \in \mathbb{R}, \quad a \neq 0 \quad (17)$$

a forma do polinómio de grau três em α , em que a, b, c e d , são constantes a determinar. Como o polinómio vai passar pelos pontos: (α_1, f_1) , (α_2, f_2) , (α_3, f_3) e (α_4, f_4) , o método dos coeficientes indeterminados origina um sistema de quatro equações lineares para calcular as constante a, b, c e d .

O mínimo do polinómio (17) verifica as condições

$$p'(\alpha) = 3a\alpha^2 + 2b\alpha + c = 0,$$

e

$$p''(\alpha) = 6a\alpha + 2b > 0,$$

cujas soluções é

$$\hat{\alpha} = \frac{-b + \sqrt{b^2 - 3ac}}{3a}. \quad (18)$$

O processo terá de ser repetido até que as condições de aceitação de um valor de $\alpha^{(k)}$ sejam verificadas. Tal como no caso da aproximação quadrática, o processo reinicia-se pela fase de procura DSC com $\alpha_0 = \hat{\alpha}$ de (18) e $\delta = \mu\delta$ ($\mu < 1$).

3.2 Algoritmo da Secção Dourada

O algoritmo da Secção Dourada tem como objectivo a procura de um intervalo que contém o mínimo de uma função objectivo unidimensional.

A diferença principal relativamente ao algoritmo de DSC reside nas condições iniciais. No algoritmo da Secção 3.1 era indicada uma estimativa inicial, α_0 , a partir da qual se procurava o intervalo que encaixasse o mínimo a estimar.

O algoritmo da Secção Dourada (SD), inicia-se com duas estimativas, os limites inferior e superior de um intervalo, suficientemente amplo, que contenha o mínimo pretendido. Neste intervalo a função deve ser unimodal, isto é, possuir um único mínimo (aquele que se pretende estimar). O esforço é então todo dirigido para a diminuição da amplitude do intervalo, calculando-se novos limites, de tal forma que o novo intervalo esteja encaixado no anterior e contenha o mínimo.

Considerando o número de ouro, $\tau = \frac{1+\sqrt{5}}{2}$, e designando por α_1 e α_2 , os valores de α que correspondem respectivamente aos limites inferior e superior do intervalo inicial, os pontos interiores, α_3 e α_4 , são calculados a partir de

$$\alpha_3 = \frac{\alpha_2 + \tau\alpha_1}{\tau+1} \quad e \quad \alpha_4 = \frac{\alpha_1 + \tau\alpha_2}{\tau+1}.$$

A escolha do subintervalo que contém o mínimo é feita de acordo com os valores da função objectivo, nos dois pontos interiores α_3 e α_4 . Dos quatro valores, α_1 , α_3 , α_4 e α_2 , será rejeitado o limite que estiver mais próximo do ponto interior (α_3 ou α_4) com maior valor da função e os restantes três serão usados para definir o intervalo que contém o mínimo para a iteração seguinte.

Na iteração que segue, apenas um novo ponto precisa de ser determinado. Novamente, a razão de ouro é a constante que ajuda a determinar a sua posição.

Este processo repete-se, calculando intervalos sucessivamente encaixados uns nos outros e que contêm o mínimo. O processo está concluído quando a amplitude do intervalo calculado for inferior a uma precisão pré-definida.

Uma explicação mais detalhada poderá ser encontrada em [1], [13] ou ainda em [14].

Após esta fase de procura do intervalo que contém o mínimo, segue-se a fase de aproximação da função por uma forma polinomial, que pode ser quadrática ou cúbica. O que foi referido em 3.1.1 e 3.1.2 aplica-se também neste caso. No caso da aproximação quadrática os três valores do intervalo seleccionado (um dos limites e os dois pontos interiores) não se encontram igualmente espaçados. A fórmula (16) exige o cálculo de um ponto extra que será definido como o ponto médio do intervalo. Para a interpolação cúbica, este ponto extra não será necessário, uma vez que os dois pontos interiores e os dois exteriores são suficientes.

3.3 Algoritmo de Davies, Swann e Campey modificado

A metodologia base deste método é igual à do algoritmo referido em 3.1. Assim, a partir de um valor inicial, α_0 , obtêm-se sucessivamente os valores

$$\alpha_1 = \alpha_0 + \delta$$

$$\alpha_2 = \alpha_1 + 2\delta$$

$$\alpha_3 = \alpha_2 + 4\delta$$

...

se a procura corre no sentido positivo, com δ positivo e o processo pára quando for definido um novo valor cujo sinal da derivada da função seja diferente do sinal da derivada no valor anterior da sucessão. Se a derivada não muda de sinal para a direita do valor inicial α_0 , a procura deve ser reiniciada, mas agora para a esquerda de α_0 .

No algoritmo DSC modificado é o sinal da derivada da função, em cada novo valor, que determina os limites inferior e superior do intervalo que contém o mínimo.

Assim, enquanto o sinal da derivada da função se mantiver, admite-se estar na presença de um intervalo de monotonia da função, e a pesquisa continua. Se o valor da derivada for negativo, conclui-se que a função é decrescente, e avança-se até encontrar um ponto cujo valor da derivada seja positivo. Se o valor da derivada for positivo, conclui-se que a função é crescente, e a procura segue no sentido negativo até se obter um ponto cujo valor da derivada é negativo.

Esta técnica garante um intervalo que contém o mínimo desejado. Para os limites do intervalo, α_1 e α_2 , conhecemos $f(\alpha_1)$, $f(\alpha_2)$, $f'(\alpha_1)$ e $f'(\alpha_2)$. Esta informação é requerida para uma aproximação cúbica da função, baseada em dois pontos, nos valores da função e da derivada (veja-se em 3.3.1). Para funções continuamente diferenciáveis este tipo de aproximação costuma ser recomendado pela sua eficiência. Além disso em 3.3.2 propõe-se outra técnica para aproximar o mínimo da função baseada na informação recolhida. Uma descrição mais detalhada do funcionamento deste algoritmo encontra-se no trabalho de Matias [14].

3.3.1 Interpolação cúbica baseada em dois pontos

Como ponto de partida desta etapa de aproximação, usam-se os valores da função objectivo f_1 e f_2 , em dois pontos dados α_1 e α_2 , bem como os respectivos valores da derivada f'_1 e f'_2 .

Assim, na sucessão daquilo que foi descrito em 3.3, falta encontrar a fórmula para o cálculo do mínimo de um polinómio de grau três, que passa pelos dois pontos, (α_1, f_1) e (α_2, f_2) cujos valores das derivadas são respectivamente f'_1 e f'_2 .

O objectivo é pois aproximar a função $f(\alpha)$ por um polinómio do terceiro grau. Seja

$$p(\alpha) = a\alpha^3 + b\alpha^2 + c\alpha + d \quad a, b, c, d \in \mathbb{R}, \quad a \neq 0$$

esse polinómio em que a , b , c e d , são constantes a determinar.

O ponto estacionário de $p(\alpha)$ é o valor de (18). Do método dos coeficientes indeterminados, resultam valores das constantes a , b , c e d , que substituídos em (18) originam

$$\hat{\alpha} = \alpha_1 + (\alpha_2 - \alpha_1) \left(1 - \frac{f'_2 + \lambda - \mu}{f'_2 - f'_1 + 2\lambda} \right) \quad (19)$$

em que

$$\mu = \frac{3(f_1 - f_2)}{\alpha_2 - \alpha_1} + f'_1 + f'_2$$

e

$$\lambda = \sqrt{\mu^2 - f'_1 f'_2} \quad (20)$$

O intervalo de interpolação $[\alpha_1, \alpha_2]$, deve verificar $f'_1 < 0$ e $f'_2 > 0$ ([21]). Este facto é de primordial importância pois

$$\mu^2 - f'_1 f'_2 \geq 0$$

e de (20) garante-se um valor real para λ . Como $f'_2 - f'_1 > 0$ o denominador de (19) nunca será zero.

3.3.2 Intersecção das tangentes

Este método, tal como o anterior, calcula uma aproximação ao mínimo de uma função, que pertence a um intervalo. Pressupõe serem conhecidos os limites inferior e superior do intervalo que contém o mínimo, α_1 e α_2 , os correspondentes valores da função, f_1 e f_2 e os valores da derivada, f'_1 e f'_2 . O algoritmo de DSC modificado, (veja-se em 3.3) fornece o intervalo pretendido, bem como toda a informação necessária, f_1 , f_2 , f'_1 e f'_2 .

O algoritmo da intersecção das tangentes, constrói duas tangentes à função, nos pontos de abcissa α_1 e α_2 , e usa o valor de $\hat{\alpha}$ que corresponde à intersecção dessas tangentes, como aproximação ao mínimo da função f .

Designemos por r e s as duas rectas que passam respectivamente por (α_1, f_1) com declive f'_1 e (α_2, f_2) com declive f'_2 . Como, do algoritmo DSC modificado, se tem

$$\alpha_1 < \alpha_2, \quad f'_1 < 0 \quad \text{e} \quad f'_2 > 0,$$

as rectas r e s intersectam-se no interior do intervalo $[\alpha_1, \alpha_2]$ ou fora dele.

Sejam

$$y = m_1x + b_1 \tag{21}$$

e

$$y = m_2x + b_2 \tag{22}$$

respectivamente as equações das rectas r e s em que $m_1 = f'_1$ e $m_2 = f'_2$. Sabendo que os pontos (α_1, f_1) e (α_2, f_2) , satisfazem as respectivas equações das rectas, determina-se facilmente os parâmetros b_1 e b_2 :

$$b_1 = f_1 - f'_1\alpha_1 \tag{23}$$

e

$$b_2 = f_2 - f'_2\alpha_2. \tag{24}$$

O ponto de intersecção das duas rectas é (substituindo (23) e (24) em (21) e (22) e igualando (21) a (22)) então

$$\hat{\alpha} = \frac{f_1 - f_2 - f'_1\alpha_1 + f'_2\alpha_2}{f'_2 - f'_1} \quad \text{com} \quad \hat{f} = f(\hat{\alpha}). \tag{25}$$

Embora o denominador de (25) nunca se anule, podem ocorrer situações em que devido a um declive muito pequeno, a intersecção se encontre fora do intervalo de interpolação. Neste caso escolhe-se o ponto médio do intervalo de interpolação como aproximação ao mínimo de f . Detalhes do correspondente algoritmo podem ser consultados em Matias [14].

3.4 Critério de Armijo

Quando não se exige que o valor de $\alpha^{(k)}$ verifique (13), podem-se implementar outros métodos mais simples, como o Critério de Armijo ou o Critério de Curry-Altman-Armijo. Estes geram valores de $\alpha^{(k)}$ e consequentemente aproximações $x^{(k)}$, que definem uma sucessão de valores da função monótona e decrescente.

A implementação do Critério de Armijo exige que sejam já conhecidos o valor $x^{(k)}$, o correspondente valor da função objectivo, $f(x^{(k)})$, e o vector direcção $d^{(k)}$. Esta deve ser de descida, para que exista um escalar $\alpha > 0$, tal que $f(x^{(k)} + \alpha d^{(k)}) < f(x^{(k)})$.

O critério de Armijo (CA) define uma sucessão de valores $\{\sigma^j\}$, $j = 0, 1, 2, \dots$, com $\sigma = 0.5$. O valor de $\alpha^{(k)}$, aceite na iteração, será o primeiro elemento desta sucessão que verifica a codição Goldstein-Armijo (veja-se na Secção 4). Esta condição tem como objectivo garantir um decréscimo significativo no valor da função, de uma iteração para a seguinte.

3.5 Critério de Curry-Altman-Armijo

No critério anterior, o primeiro valor testado para o comprimento, α , a efectuar ao longo da direcção $d^{(k)}$, é a unidade. Embora a escolha do valor inicial pareça simples, óbvia e coerente, não se tem questionado outras hipóteses, nem se tem justificado convenientemente a escolha feita. Pode-se pensar noutro valor de α mais eficiente, no sentido de que o decréscimo da função é maior e o processo iterativo torna-se mais rápido. Esta foi a ideia base que presidiu ao algoritmo de Curry-Altman-Armijo.

Os princípios de Curry e Altman determinam que a solução desejada de (13) é a menor raiz positiva da equação

$$[g(x^{(k)} + \alpha d^{(k)}) - \eta g(x^{(k)})]^T d^{(k)} = 0, \quad 0 \leq \eta < 1, \quad (26)$$

uma vez que define o ponto mais próximo de $x^{(k)}$, ao longo de $d^{(k)}$, cuja projecção do gradiente sobre a direcção é reduzida de uma quantidade η em relação à projecção do gradiente em $x^{(k)}$, (veja-se em [17]). Da linearização da equação não linear (26) obtém-se

$$\hat{\alpha} = \frac{(\eta-1)g(x^{(k)})^T d^{(k)}}{[g(z^{(k)}) - g(x^{(k)})]^T d^{(k)}}, \quad (27)$$

com $z^{(k)} = x^{(k)} + d^{(k)}$.

A escolha entre este valor de $\alpha = \hat{\alpha}$ (de (27)) e $\alpha = 1$ é feita de acordo com os valores da função que lhes correspondem. Aquele que originar o valor de f mais baixo será seleccionado.

O critério de Curry-Altman-Armijo (CCAA) define então uma sucessão $\{\alpha^j\}$, $j = 0, 1, 2, \dots$ com $\sigma = 0.5$ e aceita como $\alpha^{(k)}$ o primeiro elemento desta sucessão que verifica a condição de Goldstein-Armijo. A dedução da fórmula (27), bem como detalhes do algoritmo, podem ser encontrados em [5].

4. Resultados

Foi utilizado um computador DEC system 5000/240, com 64Mb de memória, correndo o sistema operativo ULTRIX 4.4, utilizando como única interface, um terminal de texto tipo VT100, sendo utilizado o compilador Digital Fortran77 para ULTRIX ([3]) como forma de criar o código executável do programa.

Com o objectivo de determinar o desempenho dos métodos de procura unidimensional, descritos neste artigo, fez-se um estudo comparativo que teve como base os algoritmos descritos, e que vão ser referenciados pelas siglas apresentadas na Tabela 1.

Algoritmos	Siglas	Localização	
DSC baseado em Interpolação Quadrática	DSC-IQ	3.1	3.1.1
DSC baseado em Interpolação Cúbica	DSC-IC	3.1	3.1.2
Secção Dourada baseado em Interpolação Quadrática	SD-IQ	3.2	3.1.1
Secção Dourada baseado em Interpolação Cúbica	SD-IC	3.2	3.1.2
DSC modificado baseado em Interpolação Cúbica	DSCm-IC	3.3	3.3.1
DSC modificado baseado na Intersecção das Tangentes	DSCm-IT	3.3	3.3.2
Crítério de Armijo	CA	3.4	
Crítério de Curry-Altman-Armijo	CCAA	3.5	

Tabela 1 - Siglas representativas dos diferentes métodos

Importa referir que em cada iteração, qualquer que seja o método usado na pesquisa, o valor aceite para o passo, $\alpha^{(k)}$, deve verificar certas condições.

Em virtude da dificuldade relacionada com a verificação da condição (13), considera-se indispensável utilizar outras condições que garantam que o valor calculado $\alpha^{(k)}$, origina uma nova aproximação $x^{(k+1)}$, melhor do que a anterior, $x^{(k)}$ e que o processo iterativo gerado convirja para um ponto estacionário da função ([12] e [17]). Uma primeira condição exige que a sucessão de valores da função objectivo seja monótona decrescente. Utiliza-se, para simplificar a notação, $f^{(k)}$ em vez de $f(x^{(k)})$. Para que a descida não se torne lenta demais, convém que $f^{(k+1)}$, diminua significativamente em relação a $f^{(k)}$. Este decréscimo significativo é garantido com a condição

$$C_1: \quad f^{(k+1)} \leq f^{(k)} + \mu_1 \alpha^{(k)} g(x^{(k)})^T d^{(k)} \quad (28)$$

em que $\alpha^{(k)}$ é o valor aceite do deslocamento e μ_1 é um parâmetro ($0 < \mu_1 < 1$). Quanto maior for μ_1 , mais restrita se torna a condição e o processo iterativo leva menos iterações. Esta condição é conhecida na literatura, por condição de Armijo ([8] e [9]).

Para evitar que um valor de $\alpha^{(k)}$ muito pequeno possa ser aceite por (28), é costume utilizar uma outra condição que impõe um limite ao decréscimo de f :

$$C_2: \quad f^{(k+1)} \geq f^{(k)} + \mu_2 \alpha^{(k)} g(x^{(k)})^T d^{(k)} \quad (29)$$

em que μ_2 é um parâmetro auxiliar com valores compreendidos entre μ_1 e 1.

As duas condições ((28) e (29)), quando utilizadas simultaneamente, determinam a condição de Goldstein-Armijo.

Na Tabela 2, as colunas assinaladas com ni, contêm o número de iterações necessárias até ser verificado o critério de paragem e nf representa o número de vezes que a função objectivo foi calculada ao longo do processo. O número máximo de iterações permitidas foi de 250.

Dos vinte e três problemas testados, vinte e um foram retirados do livro *More Text Examples for Nonlinear Programming Codes* de Klaus Schittkowski, [18]. A identificação usada na primeira coluna da Tabela 2 refere a numeração de [18]. Os dois últimos problemas

foram retirados de *A Class of Nonmonotone Stabilization Methods in Unconstrained Optimization*, de Grippo, Lampariello e Lucidi, [10], conhecidos por se verificar o efeito de Maratos em certas situações.

Método	DSC				SD				DSCm				CA		CCAA	
	IQ		IC		IQ		IC		IC		IT		ni	nf	ni	nf
α	ni	nf	ni	nf	ni	nf	ni	nf	ni	nf	ni	nf	ni	nf	ni	nf
201	3	43	3	44	3	52	3	52	3	44	6	111	8	11	5	11
202	5	89	5	90	6	115	6	115	5	84	9	160	7	20	100	201
203	8	163	8	162	8	137	8	137	9	171	13	248	15	20	10	21
204	4	118	5	82	5	91	7	154	5	80	9	165	9	15	9	19
206	5	96	4	75	6	110	6	110	4	68	8	145	7	15	8	17
207	6	123	6	123	6	103	6	103	6	103	6	115	22	27	11	23
208	22	469	22	510	26	453	28	637	19	422	62	1326	48	68	250*	501
209	142	1583	136	1514	70	1796	60	1575	138	1350	145	1831	49	483	250*	501
210	250	2140	250	2263	2	36	25	664	250	1831	250	5100	2	4	250*	501
211	28	431	23	349	95	2177	24	514	17	228	21	293	59	329	250*	501
212	24	379	10	154	27	513	10	189	10	139	34	485	72	417	250	501
213	6	135	12	296	11	289	7	226	8	100	8	210	9	75	F	F
240	4	77	4	79	3	52	3	562	4	71	12	222	10	15	10	21
241	28	862	18	886	47	1084	37	865	32	11095	31	420	46	327	250*	501
244	12	263	12	263	11	188	11	188	12	243	21	416	44	47	22	45
246	24	419	22	380	15	256	12	213	22	353	25	414	13	51	49	99
247	21	435	21	436	24	409	23	392	21	406	250	5492	250	279	49	99
256	22	373	30	537	42	685	14	298	21	994	25	448	57	213	30	61
257	14	244	21	315	57	1149	45	988	16	284	19	222	27	223	227	455
258	66	1593	13	260	13	253	13	244	46	1169	11	186	19	54	13	27
259	76	1158	32	486	34	699	58	1149	78	3636	139	1905	109	740	250	501
M ₁	63	986	49	807	15	365	14	283	61	875	59	759	44	274	250	501
M ₂	250	2635	237	2715	48	1290	245	6104	250	2361	162	2289	228	2525	250	501

Tabela 2 - Resultados dos problemas

A primeira coluna da Tabela 2 referencia o problema, as quatro colunas seguintes contêm os resultados referentes a implementações do método baseado na procura DSC nas aproximações quadrática e cúbica. Destas, a aproximação cúbica foi mais rápida em 48% dos casos e a quadrática em 39%. As quatro colunas seguintes referem-se à procura da Secção Dourada com aproximações quadrática e cúbica. Esta última é novamente mais eficiente, pois dos vinte e três problemas testados resultaram 52% de casos com menor número de iterações ou de cálculos da função. A aproximação quadrática, por sua vez foi mais eficiente em 22% dos casos. A aproximação cúbica com a forma DSC foi mais eficiente (em 52% dos problemas) do que com a procura SD. Da procura DSC modificada, conjugada com a interpolação cúbica ou a intersecção das tangentes como técnicas de aproximação, resultaram os valores das colunas 10, 11, 12 e 13 da Tabela 2. Com a interpolação cúbica obtiveram-se os melhores resultados: convergências mais rápidas (nalguns casos diferenças bem significativas) em 70% dos casos. Apenas em 17% a intersecção das tangentes foi melhor do que a interpolação cúbica. Em dois problemas o DSCm-IC precisou de menos iterações do que o DSCm-IT mas foi necessário quase o dobro de

cálculos da função. Se compararmos os dois algoritmos baseados em interpolação cúbica, (nas colunas 4 e 5) com uma procura básica DSC, com as colunas 10 e 11, obtidas com a procura DSC modificada, verifica-se que em 30% dos casos o primeiro algoritmo foi mais eficiente pois exigiu menos iterações. Em 17% dos casos, o DSCm-IC foi mais eficiente. Em 10 dos 23 problemas, a interpolação cúbica baseada na procura DSCm exigiu ou menos iterações ou menos cálculos da função, no entanto, a diferença não é considerada significativa, uma vez que o algoritmo também utiliza cálculos da derivada.

As últimas quatro colunas representam os valores obtidos pelos critérios de Armijo e de Curry-Altman-Armijo. O primeiro é mais rápido em 70% dos casos e o segundo em 30% dos casos.

Comparando o critério de Armijo com o método de procura DSC baseado em interpolação quadrática (talvez os mais conhecidos e usados) verifica-se que em 43% dos casos o DSC-IQ leva menos iterações, no entanto, considerando o número de cálculos da função, apenas num caso o número é inferior ao do critério de Armijo. Os critérios de Armijo e de Curry-Altman-Armijo, menos exigentes em termos de esforço computacional necessário em cada iteração, apresentam valores de nf muito mais baixos do que qualquer dos outros métodos. Nalguns casos chegam a atingir valores seis vezes mais baixos do que nos métodos DSC, SD e DSC modificado.

O menor esforço computacional conjugado com uma maior eficiência, torna estes algoritmos muito usados e seleccionados para as mais diversas aplicações.

Os asteriscos na Tabela 2 representam casos em que se obteve convergência embora o critério de paragem não tivesse sido verificado antes de atingir as 250 iterações. Os casos em que o limite de iterações, 250, está indicado na Tabela 2 referem sucessos e a paragem verificou-se nessa iteração. Apenas num caso não se verificou convergência e está assinalada com um F.

5. Conclusões

Como conclusão, e no que concerne aos algoritmos de pesquisa unidimensional, é importante realçar a eficiência em termos do número de cálculos de valores da função dos critérios de Armijo e Curry-Altman-Armijo, quando combinados com direcções Quasi-Newton. O algoritmo que implementa o critério de Armijo origina implementações bem sucedidas em 100% dos casos. O critério de Curry-Altman-Armijo, embora tenha uma estrutura algorítmica um pouco mais complicada do que a de Armijo, deu origem a implementações que convergiram, em menos de 250 iterações, para a solução pretendida, em 74% dos casos.

A experiência adquirida com a implementação dos oito algoritmos diferentes, mostra que a pesquisa unidimensional baseada nos critérios de Armijo e de Curry-Altman-Armijo é, no global, mais eficiente quando comparada com os restantes algoritmos, mas podem falhar em certas situações.

A robustez do algoritmo CCAA anda, nesta altura, pelos 74% quando implementado com direcções do tipo Quasi-Newton e ronda os 96% quando implementado com direcções do tipo Newton. Este valor da robustez foi verificado no trabalho de Fernandes [5], superior à verificada com o algoritmo do critério de Armijo, 88%. Este facto leva-nos a concluir que o critério de CCAA é mais adequado quando conjugado com métodos de convergência quadrática. Embora estas últimas conclusões tenham sido tiradas no âmbito da resolução de problemas com restrições nas variáveis, a validade mantém-se uma vez que a estratégia utilizada na sua resolução determina uma optimização sem restrições. A função objectivo resultante, é na maior parte dos casos, ainda mais complexa e difícil de minimizar que as funções dos problemas tratados neste artigo.

Uma razão para este valor mais baixo da robustez parece ser a que se relaciona com o facto do método Quasi-Newton, quando baseado na fórmula de actualização de DFP (veja-se a Secção 2), ser globalmente convergente com funções convexas, apenas quando conjugado com uma procura unidimensional exacta. Recomenda-se a leitura do trabalho de Nocedal [16] sobre este assunto.

Face a estes valores, pretende-se, num futuro muito próximo, analisar a estrutura algorítmica do critério CCAA e modificá-la, por forma a garantir mais sucessos com métodos de convergência superlinear. Estas alterações devem ser realizadas de modo a não baixar a eficiência do algoritmo.

Referências

- [1] Abdy, P.R. e Dempster, M.A.H., *Introduction to Optimization Methods*, Chapman and Hall, London (1974).
- [2] Celis, M.R., Dennis, J.E. e Tapia, R.A., *A Trust Region Strategy for Nonlinear Equality Constrained Optimization*, em P.T.Boogs, R.H.Byrd e R.B.Schnabel, editores, *Numerical Optimization* (1984) 71-82, SIAM, Philadelphia.
- [3] Cered, R.L.D. e Maldonado, J.C., *Introdução ao Fortran 77 para Microcomputadores*, McGraw-Hill, São Paulo, Lisboa (1987).
- [4] Fernandes, E.M.G.P., *Newton Based Exact Penalty Techniques for Nonlinear Optimization with Constraints*, em U.Derigs, A.Bachem e A.Drexel, editores, *Operations Research Proceedings* (1994) 39-44, Springer-Verlag, Berlin.
- [5] Fernandes, E.M.G.P., *A Curry-Altman-Armijo Line Search Algorithm which uses Newton Based Descent Directions*, em *Proceedings of Applied Mathematical Programming and Modelling* (1995) 27-30, London.
- [6] Fernandes, E.M.G.P., *Computação Numérica*, Serviços de Reprografia e Publicações da Universidade do Minho, Braga, Portugal, 2ª edição (1998).
- [7] Fletcher, R., *Practical Methods of Optimization*, John Wiley & Sons, New York, 2ª edição (1991).
- [8] Gill, P.E. e Murray, W., *Safeguard Steplength Algorithms for Optimization using Descent Directions*, Technical Report 37, National Physical Laboratory, Division of Numerical Analysis and Computing, Teddington (1974).
- [9] Gill, P.E., Murray, W. e Wright, M.H., *Practical Optimization*, Academic Press, London (1981).
- [10] Grippo, L., Lampariello, F. e Lucidi, S., *A Class of Nonmonotone Stabilization Methods in Unconstrained Optimization*, *Numerische Mathematik* 59 (1991) 779-805.
- [11] Dennis Jr, J.E. e Moré, J.J., *A Characterization of Superlinear Convergence and its Application to Quasi-Newton Methods*, *Mathematical of Computation* 28 (1974) 549-560.
- [12] Dennis Jr, J.E. e Schnabel, R.B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, New Jersey (1983).
- [13] Luenberger, D.G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company (1973).
- [14] Matias, J.L.H., *Algoritmos de Pesquisa Unidimensional na Optimização de Funções, Estudo Comparativo*, Tese de Mestrado, Escola de Engenharia da Universidade do Minho, Braga (1996).

- [15] Monteiro, M.T. e Fernandes, E.M.G.P., *Uma Procura Curvilínia na Otimização Não Linear com Restrições*, em Actas do III Congresso Galego de Estatística e Investigación de Operacións (1997) 269-276, Lugo-Spain.
- [16] Nocedal, J. *Theory of Algorithms for Unconstrained Optimization*, em Acta Numerica (1992) 199-242, Cambridge University Press.
- [17] Ortega, J.M. e Rheinbold, *Iterative Methods of Nonlinear Equations in Several Variables*, Academic Press, London (1970).
- [18] Schittkowski, K., *More Text Examples for Nonlinear Programming Codes*, Economics and Mathematical Systems, Springer-Verlag, Berlin (1987).
- [19] Soares, J.L.C., *Um Estudo sobre o Método de Newton para Otimização sem Restrições*, Investigação Operacional 14 (1994) 233-258.
- [20] Vicente, L.N., *A Comparison Between Line Searches and Trust Regions for Nonlinear Optimization*, Investigação Operacional 16 (1996) 172-179.
- [21] Wolfe, M.A., *Numerical Methods for Unconstrained Optimization*, Van Nostrand Reinhold Company, New York (1978).

MULTICAST ROUTING TREE OPTIMISATION

G. Di Fatta

G. Lo Re

CE.R.E. - Centro studi sulle Reti di Elaboratori
C.N.R. - Consiglio Nazionale delle Ricerche
Viale delle Scienze - Parco d'Orleans
90128 Palermo - Italy

Abstract

High bandwidth networks allow new distributed applications which involve all the hosts of a target group. These applications send messages to all members of the group, and such multipoint communications require techniques which must efficiently use network resources. An effective method of multicast routing is achieved by using a spanning tree which connects all target nodes in the network. An important objective is to minimize the overall network cost of the spanning tree. This is known as the Steiner Tree Problem in networks, and it is proved to be NP-complete. Many heuristics have been proposed in the past years, which are capable of isolating sub-optimal solutions in polynomial times. In this paper we present a new method which is used to obtain more accurate approximations of the optimal solutions. An algorithm is proposed which is capable of improving the solutions obtained by well known heuristics, by means of a stirring process of the nodes in solution trees. Solutions obtained by the best available heuristics are used as starting points. An enumerative method is also used as comparison term in the experimental analysis which demonstrates the goodness of the method discussed in this paper.

Keywords

Bandwidth networks, networks, spanning trees.

I. Introduction

Multicasting is the simultaneous transmission of data to multiple destinations. The simplest solution for the multicast routing involves the construction of a delivery tree whose effectiveness measure is the sum of the costs on its edges.

The construction of such minimal multicast tree entails the resolution of a difficult optimization problem, universally known as Steiner Tree Problem. This problem has been proved to be NP-complete and many heuristics have been proposed capable of isolating sub-optimal solutions in a polynomial time. Among these, in this paper we will consider the Shortest Path Heuristic (SPH) [7], the Distance Network Heuristic (DNH) [9], the Kruskal-based SPH (KSPH) [5], the Average Distance Heuristic (ADH) [6] and the repetitive SPH (SPH-Z) [10].

The solution trees produced by these heuristics can be even further improved, using some post processing techniques.

This paper proposes a new technique that is capable of optimising the solutions produced by classical heuristics by means of a total repositioning of the tree edges. Another advantage of this method consists in obtaining solutions which are competitive with the ones obtained by more complex heuristics.

The terminology and notation adopted are described in Section II, and some classical heuristics are summarised whose results will be used as bases for the application of our method. The heuristic performance guarantees and complexities are also presented.

Section III introduces the new algorithm which improves the trees generated by classical heuristics.

The final part of the paper (section IV) is devoted to experimental analysis, based on the results of a wide range of tests.

II. Steiner Tree Heuristics

In this section a short summary of previous Steiner Tree Heuristic is presented, and some basic notations and terms are therefore introduced. Given a graph $G = (V, A)$ with node set V and edge set A , we will use $Z \subset V$ to indicate the subset of multicast nodes and $S = V - Z$ to indicate non-terminal nodes.

Many heuristics proposed in the past years have been capable of isolating sub-optimal solutions involving polynomial worst-case time complexities. Among these, the Minimum or Shortest Path Heuristic (MPH or SPH), the Distance Network Heuristic (DNH), the K-Shortest Path Heuristic (K-SPH), the Average Distance Heuristic (ADH), the repetitive Shortest Path Heuristic (SPH-Z) and so on.

Some of these heuristic which we used as bases for our algorithm are briefly described here.

The DNH builds the distance network K_z introduced by Z . It constructs the minimum spanning tree (MST) on the K_z network and replaces the virtual links with the real paths (the nodes and links of the initial network), thus obtaining G_z , a subgraph of the initial network. It then computes the MST on G_z and finally prunes all the S -vertices of degree one.

The SPH builds a subtree of G in an incremental fashion: it starts off by selecting an arbitrary node among the terminal nodes (typically the source node) and then progressively adds the terminal node nearest to the tree, including the nodes and edges of the connecting path.

The K-SPH is an improvement of the SPH algorithm. It builds a forest of subtrees joining together the closest Z -nodes or subtrees until a single solution tree has been obtained.

A straightforward improvement of the SPH is to run it for every possible choice of the fixed Z -vertex, and then take the best solution found. We refer to this heuristic as the repetitive Shortest Path Heuristic (SPH-Z).

ADH is a generalisation of K-SPH. It repeatedly connects the two closest multicast group member components through the most central node. ADH terminates when a single tree remains, spanning all the Z -nodes.

The competitiveness of these heuristics can be measured in terms of the ratio between the cost of the solutions they are able to identify, and the cost of the optimal solutions. Many algorithms capable of identifying the optimal solution tree have been proposed in the past. All of them are characterised by an exponential complexity. Among these, the Spanning Tree Enumeration Algorithm [9], whose solutions we have considered as comparison terms, has the complexity shown in the table below.

Table 1 summarises the complexities and performances for the algorithms discussed above.

In the table, m is the number of the terminals, and n is the number of nodes in the network. It should be noticed that all the heuristics discussed above have polynomial complexity in $O(mn^2)$. The repetitive heuristic SPH-Z is the most complex one with $O(mn^2+nm^3)$. In literature ADH is reported to be the best performing one among single step heuristics, but its complexity $O(n^3)$ is an order higher. Moreover it is worse than the repetitive SPH-Z. Other more complex repetitive heuristics have been proposed in [10]. However their worst-case time complexities are non comparable with the ones of the single step methods.

Algorithm	Performance Guarantee	Complexity
DNH	2	$O(mn^2)$
SPH	2	$O(mn^2)$
K-SPH	2	$O(mn^2)$
ADH	2	$O(n^3)$
SPH-Z	2	$O(mn^2+nm^3)$
STEA	1	$O(m^2 2^{(n-m)} + n^3)$

Table 1 - Performance Guarantee and Complexity

III. Post Processing Techniques

As noted by Rayward-Smith and Clare [6], finding the Steiner Minimal Tree in a graph G for a set Z of terminal nodes is equivalent to finding the Minimum Spanning Tree in the sub-graph of G induced by $Z \cup S'$, where S' ($S' \subseteq S$) is the set of Steiner vertices, i.e. those non-terminal nodes of the Steiner Minimal Tree.

Solutions found by all the classical heuristics could be improved using a post-processing technique introduced by Rayward-Smith and Clare (RSC). This method finds the graph induced by the nodes in the solution tree, derives the minimum spanning tree of the induced graph, and prunes this tree of non-multicast leaves. The resulting tree is an improved solution.

This method does not involve the insertion of any additional node of the original graph, but on the sub-network induced by the nodes of the solution tree re-arranges nodes and links.

In the following, we will introduce a new post-processing technique which we will refer as "Stirring" algorithm.

The idea underlying the Stirring algorithm is that alternative and lower cost paths connecting nodes of the tree backbone can be individuate at the end of the heuristic execution. The tree construction process of most heuristics involves the addition of some Z -node z_i to a

transitory sub-tree T_i . Nevertheless, the minimum path between the selected node z_i and the successive sub-trees T_j (with $j > i$) does not remain constant during the tree growing process. This is because the addition of further nodes into the tree could reduce the cost of the connecting path of a previously inserted node. For this reason, we considered the possibility of recalculating the "grafting" point of some node of the tree at the end of the tree construction. By the term 'grafting point' of a node n_i in a tree, we mean the closest ancestor of n_i which has at least two children.

For any node n_i , we check the existence of a node n_k in the tree with $d(n_i, n_k) < d(n_i, n_a)$ where n_a is the current grafting point and n_k does not belong to the path $\langle n_i, n_a \rangle$, and is not a descendant of n_i . If such a node n_k is found in the tree, the path $\langle n_i, n_a \rangle$ is replaced by the $\langle n_i, n_k \rangle$ one. This procedure is executed for each node which has at least two children and for the leaves of the tree.

Path replacement may provoke the insertion of new nodes into the tree. Increasing the subset of nodes extends the solution space explored by heuristics. The Stirring process is repeated until no further improvement can be obtained in the solution tree.

The previous methods (RSC and Stirring) are not mutually exclusive. The RSC can be applied after the Stirring algorithm execution.

IIIA. Pseudo-code and complexity analysis

(Stirring post-processing technique)

While some change occurs

For each node n in Tree with $\text{degree}(n) \geq 3$ or $= 1$

Search a better grafting point for n ;

If a better grafting point exists

Replace the path in the tree,

End while

(RSC post-processing technique)

Let G_z be the subgraph of G induced by $Z \cup S_T$

Compute T as the MST of the subgraph G_z

Delete from T all S -vertices of degree 1 (one at a time)

The resulting tree is the suboptimal solution.

As described in the pseudo-code, the Stirring algorithm searches for a better grafting point among the tree nodes (at most n). This is done for each leaf and for each grafting point (degree higher or equal to three) in the tree (at most $2m$ nodes). The whole process is repeated until no further changes occur (in the worst case n times). This results in $O(n^2m)$ complexity of the algorithm.

The Rayward-Smith & Clare post processing algorithm entails a worst-case time complexity which is the one of the minimum spanning tree algorithm, i.e. $O(n^2)$.

IIIB. Examples

In the following figures we show the results of the execution of the RSC and Stirring post processing techniques when they are applied to the network of Figure 1a.

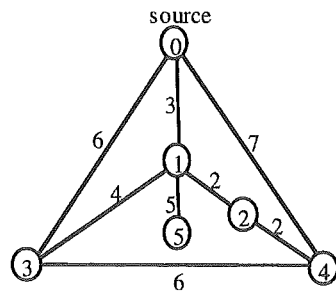


Figure 1a - 6-node network

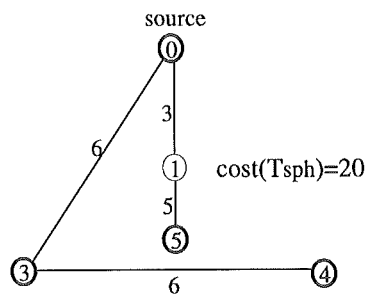


Figure 1b - SPH solution tree

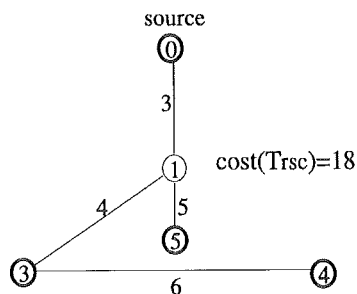


Figure 1c - RSC solution tree

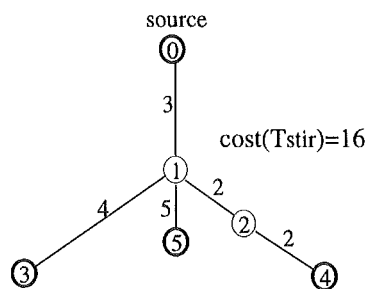


Figure 1d - Stirring solution tree

In Figure 1b we show the solution tree as obtained by the SPH heuristic. The SPH solution tree has been adopted as starting point of both post processing techniques.

Figure 1c represents the tree isolated by the RSC method. The minimum spanning tree obtained over the graph induced by the nodes of the SPH tree entails a cost reduction (-2) due to the adoption of the edge <1,3> in place of the edge <0,3>.

Figure 1d is the solution determined by the Stirring method. Its total cost reduction (-4) is obtained by means of the detection of a better grafting point (node 1) for the nodes 3 and 4. In particular the node 4 is connected by means of the path <1,4> in place of the edge <3,4>. It should be noticed that this path replacement has involved the addition of a new node (node 2) to the SPH solution tree.

IV. Experimental Analysis

To generate the network for our experiments we adopted the Tiers [4] tool. It is capable of randomly generating graphs whose characteristics are those of real computer networks. It is important to ensure that the network models being used to test the algorithm bear some resemblance to the real networks which generally are very sparse graphs, i.e. with a number of edges lower than twice the number of nodes.

The Tiers tools uses the classical Waxman model for generating graphs, i.e. the nodes in the network are distributed at random across a Cartesian co-ordinates grid. The generic nodes u and v are connected by a link which is set according to the following probability law:

$$P(u,v) = \beta e^{\frac{-d(u,v)}{L\alpha}} \quad (1)$$

where $d(u,v)$ is the Euclidean distance between u and v , L is the maximum distance between two nodes, and α and β are parameters in the range $0 < \alpha, \beta \leq 1$. A large value of α increases the number of connections to nodes further away, whilst a large value of β increases the number of edges from each node.

Furthermore, a series of modifications to the Waxman model make the generated networks more similar to real networks by using different hierarchies of nodes. Parameters are required to define the topological characteristics, such as the number of WAN, MAN and LAN, the number of nodes in each subnetwork, and the different connection degrees [3]. We generated different network sets to test our algorithms. These sets were characterised by different sizes (200, 500 and 1000 network nodes), and different link redundancies. A high redundancy degree involves the presence of more complex solutions. The chosen multicast group sizes were 5%, 10% and 20% of the graph nodes.

IVA. Enumeration method

The knowledge gained about the optimal solution allows us to make a precise comparison of the different heuristics. This is necessary to understand how often the classical heuristics fail in terms of optimal solution detection, meaning the cases where it is possible to obtain better solutions. To this end we implemented the Spanning Tree Enumeration algorithm [9].

Furthermore, because the NP-complete nature of the Steiner Minimal Tree problem makes the optimal solution search a computationally heavy task, we used the reduction rules, as

described in [9], in order to increase the size of the explored networks. By means of these rules a particular instance of the STP can often be reduced to a smaller one by examining local properties of the networks.

Networks with two hundred nodes and ten terminal nodes are reduced to networks with 50 nodes and 8 terminals on average. They are the ones for which we were able to find the optimal solution in a few minutes.

IVB. Experimental results

The first experiment was devoted to determining the actual performance gain introduced by the Stirring process, as described in section III.

We selected networks and terminal sets for which we were able to find the optimal solution in a short time. We carried out a series of tests over three hundred networks, each with 200 nodes and 10 terminal nodes.

As already mentioned, among the single-pass heuristics of the literature, we chose and implemented the following three: the Distance Network Heuristic, the Shortest Path Heuristic and the Kruskal-based Shortest Path Heuristic. Furthermore, we implemented the repetitive variant of the Shortest Path Heuristic (SPH-Z). It should be noticed that some heuristic descriptions include the RSC post-processing technique as final step. We implemented each heuristic in its simplest form, to put in evidence the actual gain produced by post-processing techniques.

Over the four heuristics we applied both the post processing techniques, that means only the RSC, and the Stirring method with the final RSC. For the SPH-Z, after each repetitive call of the SPH, the algorithm involves the post-processing execution.

In the chart of Figure 2, we show the number of optimal solutions found by each heuristic when executed with and without the post processing methods.

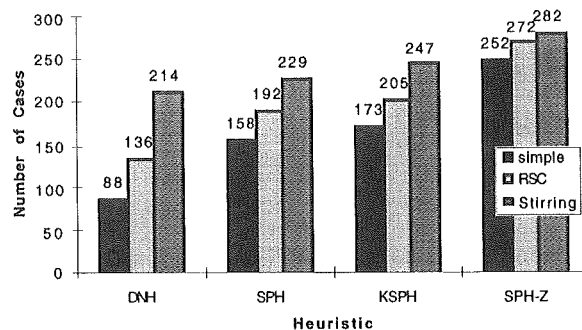


Figure 2 - Optimal solutions in 200 node networks

As shown in the previous chart, both the post processing methods are capable of sensibly improving the number of optimal solutions found. However the Stirring method is able to improve the performance of the heuristics twice the RSC technique. The SPH-Z with Stirring isolates 282 optimal solutions over the whole sample of 300 networks, i.e. the 94% of cases.

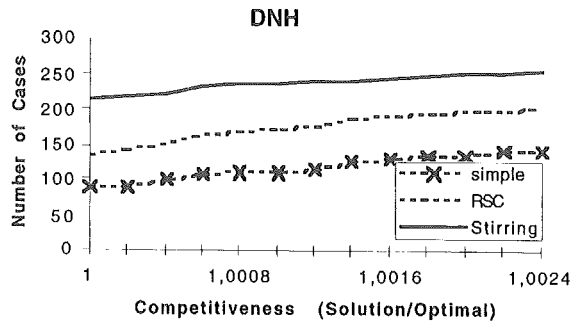


Figure 3a - Competitiveness in 200 node networks

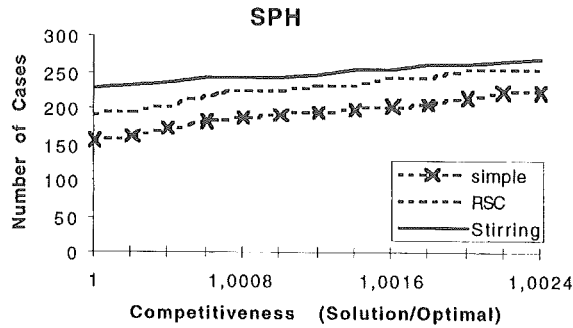


Figure 3b - Competitiveness in 200 node networks

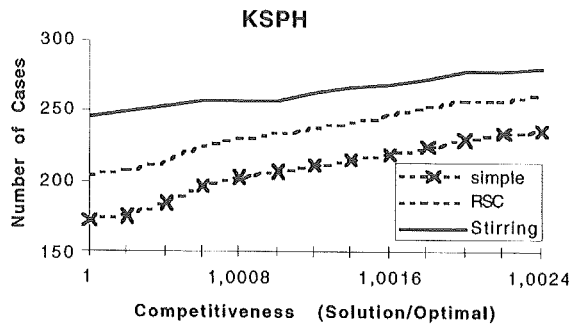


Figure 3c - Competitiveness in 200 node networks

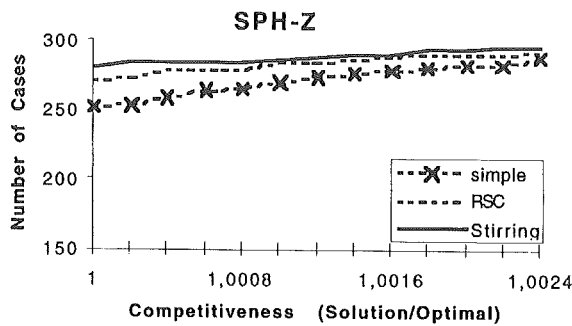


Figure 3d - Competitiveness in 200 node networks

In regard to the same experiment, the following figures show the competitiveness, determined as the ratio between the costs of trees produced by heuristics and the optimal ones. Each figure is related to a single heuristic when executed with and without the post processing methods. In all the cases, the Stirring effectiveness is confirmed also in the isolation of the sub-optimal solutions.

In the last diagram (the SPH-Z case in Fig.3d) the effectiveness of the Stirring is less significant, because the repetitive approach is able to isolate a very high number of optimal solutions in these networks with a small number of nodes.

It should be noticed that Figures 3a and 3b differ from Figures 3c and 3d in the scale adopted for the y-axis.

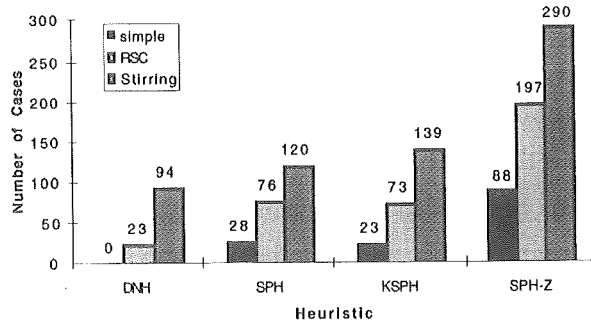


Figure 4a - Sub-optimal solutions in 500 node networks

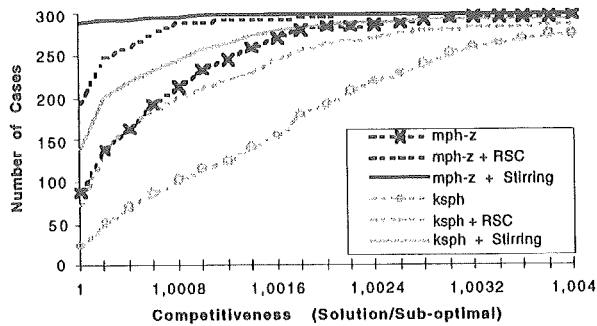


Figure 4b - Competitiveness in 500 node networks

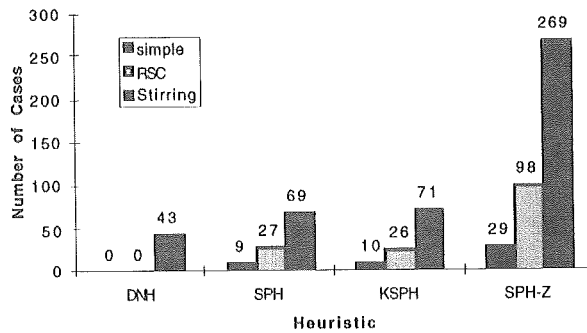


Figure 5a - Sub-optimal solutions in 1000 node networks

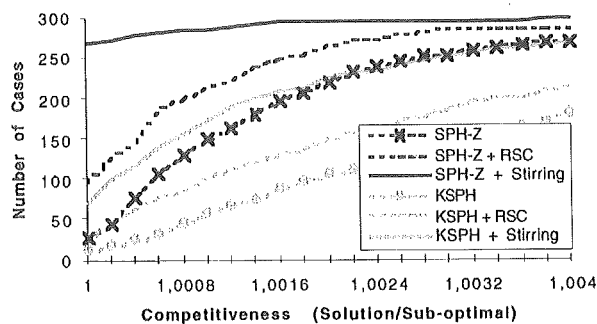


Figure 5b - Competitiveness in 1000 node networks

The experiments have been repeated for bigger network sizes, 500-node (Figs.4) and 1000-node (Figs.5) networks.

The bigger network sizes do not allow the optimal solution to be isolated in a short time. Consequently, competitiveness is determined as the ratio between heuristic and sub-optimal tree cost.

The sub-optimal solution is determined as the best solution found by any of the heuristics implemented.

The increase in the problem size determines a performance improvement of the Stirring method in comparison with the other two algorithms. In all the network samples, the Stirring isolate better solutions in more cases than the traditional ones. In particular, the gaps between Stirring and RSC become more significant when the problem size increases.

Table 2 summarises and compares all the results obtained by the heuristics over all the 1000-node networks. Each row shows how often one of the heuristics outperformed each of the other ones.

		DNH			SPH			KSPH			SPH-Z		
		simp	RSC	STR	simp	RSC	STR	simp	RSC	STR	simp	RSC	STR
DNH	simple	-	0	0	13	8	2	13	8	4	0	0	0
	RSC	298	-	6	71	32	14	69	33	8	22	3	2
	STR	300	281	-	237	175	66	219	166	76	132	64	12
SPH	simple	287	229	52	-	0	0	128	60	25	0	0	0
	RSC	292	262	91	225	-	2	211	102	40	74	0	0
	STR	298	279	157	274	203	-	242	191	72	145	74	0
KSPH	simple	287	229	77	120	63	42	-	0	0	19	5	3
	RSC	292	264	124	218	112	68	218	-	2	86	12	7
	STR	296	290	196	264	225	124	271	205	-	168	99	16
SPH-Z	simple	300	278	158	270	198	131	260	194	114	-	0	0
	RSC	300	294	208	285	253	177	280	253	162	228	-	3
	STR	300	298	255	289	271	229	286	272	227	267	267	-

Table 2 - 1000-node networks: all-to-all comparisons

The grey cells refer to the comparison between the RSC post-processing technique and the Stirring one. Our technique overcomes the other one in almost all the cases. For instance in the SPH-Z case, as shown in the last row and in the last column, among 300 networks the Stirring isolates a better solution in 267 cases, RSC only in 3, and in the remaining 30 cases they find the same solutions.

All the experiments were carried out on two UNIX workstations, a Sun Ultra30 with the SunOS 5.6 (250MHz and 128MB of RAM), and an Intel Pentium Pro based workstation with LINUX OS (233MHz and 64MB of RAM).

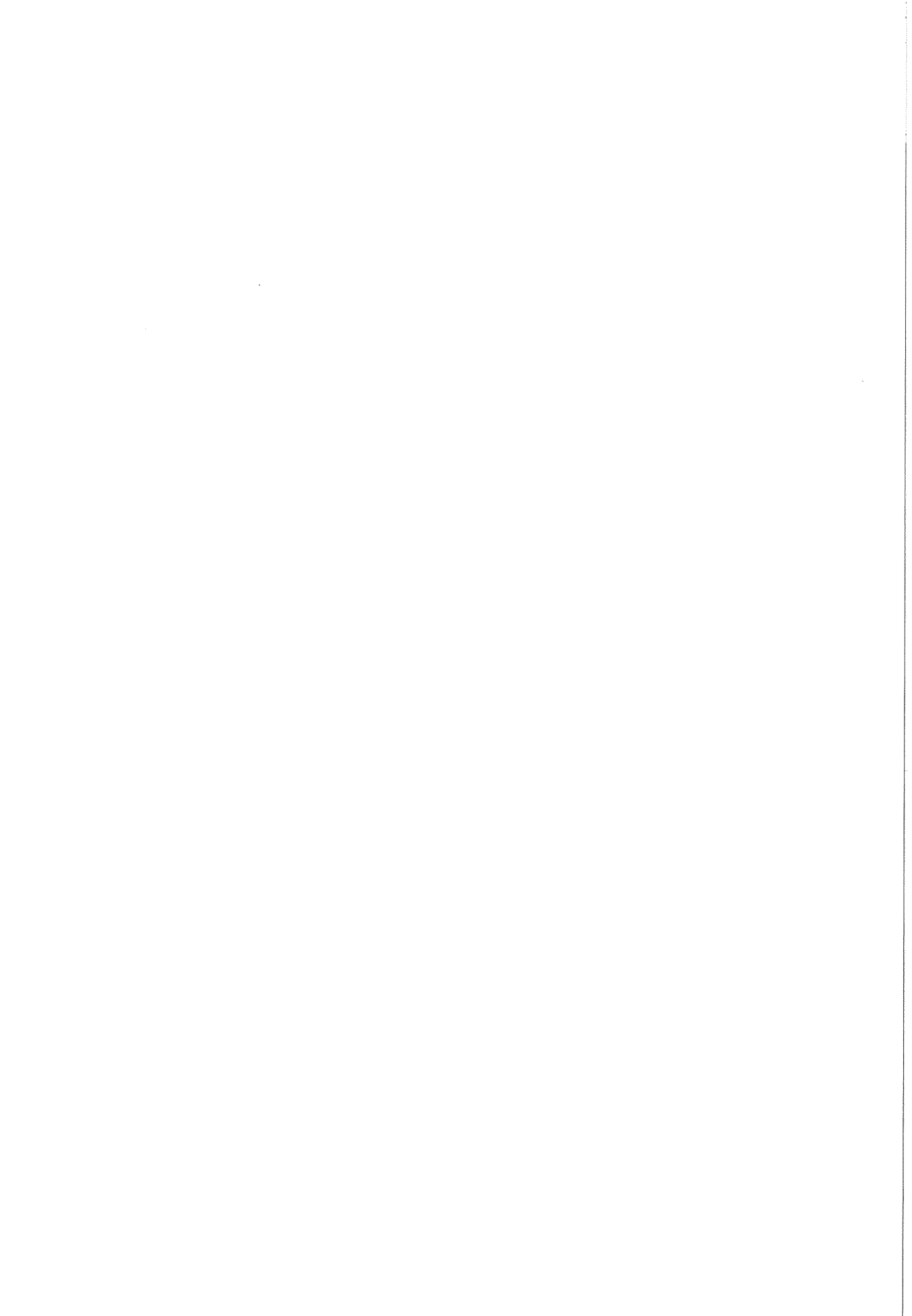
V. Conclusions

The experimental results confirmed the effectiveness of the Stirring post-processing technique.

The results also showed that the solution found by single step methods with the final Stirring are competitive with the solutions found by the repetitive SPH-Z. This means that algorithms whose worst-case time complexities are polynomial in $O(mn^2)$, using the post processing method which is also polynomial in $O(mn^2)$, are capable of reaching and in most cases improving solutions found by the repetitive algorithm. Furthermore, the improvement obtained in the SPH-Z with the Stirring also means that the solutions isolated by repetitive heuristics are sensible to the Stirring post-processing. We limited our analysis to $O(n^2)$ heuristics because of their possible usage into network protocols. How effective the post-processing method could be on the higher order heuristics has still to be investigated.

References

- [1] Bharath-Kumar, K. and Jaffe, Routing to Multiple Destinations in Computer Networks, IEEE Trans. on Commun. 31 (1983) 343-351.
- [2] Dijkstra, E.W., A Note on Two Problems in Connection with Graphs, Numer. Math. 1 (1959) 269-271.
- [3] Doar, M.B. and Leslie, I., How Bad is naive multicast routing?, Proc. of IEEE Infocom. (1993) 82-89.
- [4] Doar, M.B., TIERS User Manual, v. 1.1.
- [5] Kruskal, J., On the Shortest Spanning Subtree of a Graph and the Travelling Salesman Problem, Proc. Amer. Math. Soc. 7 (1956) 48-50.
- [6] Rayward-Smith, V.J. and Clare, A., Finding Steiner Vertices, Networks 16 (1986) 283-294.
- [7] Takahashi, H. and Matsuyama, A., An Approximate Solution for the Steiner Problem in Graph, Math. Japonica 24 (1980) 573-577.
- [8] Waxman, B., Routing of Multipoint Connections, Journal on Selected Areas in Communications 6 (1988) 1617-1622.
- [9] Winter, P., Steiner Problem in Networks: A Survey, Networks 17 (1987) 129-167.
- [10] Winter, J. and Smith J.M., Path-Distance Heuristics for the Steiner Problem in Undirected Networks, Algorithmica 7 (1992) 308-327.



CHARACTERISTICS AND APPLICATIONS OF RANDOM SEARCH METHODS FOR MATHEMATICAL PROGRAMMING

Santosh Kumar

Dept. of Applied Mathematics
National University of Science and Technology
P.O.Box 346, Bulawayo - Zimbabwe

Fernando Scarmozzino

Dept. of Computer and Mathematical Sciences
Victoria University of Technology
P.O.Box 14428, MMC, Melbourne - Australia

Abstract

In this paper we discuss the development of the random search method for mathematical programming problems. We present the characteristics of the random search approach for mathematical programming problems, and illustrate these characteristics in deriving a random search solution for a non-convex mathematical programming problem. We demonstrate its use in the ranking of the bases of a linear programming problem. A single machine scheduling situation is given and a random search method is employed in approximating a good solution. We also outline a real-life situation where an analytical solution may be difficult to obtain, but for which a random search may be useful.

Keywords

Mathematical programming, random search algorithms, single machine scheduling.

1. Introduction

Random search methods have been developed to obtain approximately near-optimal solutions for multi-dimensional functions. Random search methods appear to have been first applied in the area of experimental design. The motivating factor was the need to find methods of solving maximum-seeking problems with a large number of factors, where the steps involved in solving such problems by traditional methods proved excessive.

According to Brooks [2] two things must be defined *a priori* by the experimenter:

(i) A *definition of a successful solution* must be given. In experimental designs this is difficult to pre-determine but for mathematical programming this is far more clear-cut. For a maximization problem, this would mean obtaining an objective function value at least as large as its previous value; and (ii) An *uncertainty (probability) value of the success of the experiment*. With any such method, the experimenter needs to decide on what assurance of

success is required. This is given by *PROB*, the probability of success. Although, ideally, a large value of *PROB* should be used, this needs to be balanced against the number of steps required for a large *PROB*. The larger the value of *PROB* the greater the number of trials required.

Brooks also developed an elementary formula for estimating the number of trials required to find a solution for a given probability of assurance.

The majority of the work using random search methods has dealt with global optimization problems. A selection of such applications can be found in [4, 11, 13, 15].

The random search method of Brooks has been modified by Huynh et al. [18, 19, 21, 24] to find solutions for the following mathematical programming problems:

- (i) pure integer programming,
- (ii) mixed integer programming,
- (iii) hyperbolic integer programming, and
- (iv) extreme point mathematical programming.

Mathematical statements of these problems are presented in Section 2. The se methods have certain common elements which, in Section 3, are presented as characteristics of the random search methods. In Section 4 we define the terms and notation which will be used in the random search method. In Section 5 we outline a non-convex mathematical programming problem and develop a random search solution method for this problem. We also apply the random search technique [24] to rank the extreme points of a linear programming problem in Section 6. Finally, in Section 7, we discuss two problems, a facility location situation and a single machine scheduling problem which tend to be difficult to solve analytically. We outline how the random search method could be used to improve the solution strategy for the former and demonstrate its use via an example for the latter.

2. Mathematical Statement of Problems Solved Using Random Search Methods

In this section we present, for completeness, the mathematical statement of the above mathematical programming problems, for which random search methods and algorithms have been developed by Huynh et al. [18, 19, 21, 24].

2.1 Pure Integer Programming Problem

Mathematically the pure integer programming problem is defined as follows:

$$\begin{aligned} &\text{Maximize } z = CX \quad \text{subject to } AX \leq b \quad X \geq 0 \quad \text{and } X \in \mathbb{Z}^{n \times 1} \\ &\text{where } L \leq X \leq U, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^{m \times 1}, \quad C \in \mathbb{R}^{1 \times n}, \end{aligned}$$

and L and U are the integer lower and upper bounds for X such that $L, U \in \mathbb{Z}^{n \times 1}$.

2.2 Mixed Integer Programming Problem

$$\begin{aligned} &\text{Maximize } z = C_1X + C_2Y \quad \text{subject to } A_1X + A_2Y \leq b \quad X \geq 0 \quad Y \geq 0 \\ &C_1 \in \mathbb{R}^{1 \times n_1}, \quad C_2 \in \mathbb{R}^{1 \times n_2}, \quad n_1 + n_2 = n, \quad X \in \mathbb{R}^{n_1 \times 1}, \quad \text{and } Y \in \mathbb{Z}^{n_2 \times 1}. \end{aligned}$$

2.3 Extreme Point Mathematical Programming Problem

Maximize $z = CX$ subject to $AX \leq b$ where X is an extreme point of the set H ,
 $H = \{X: DX \leq e \text{ and } X \geq 0\}$, $A \in \mathbb{R}^{q \times n}$, $b \in \mathbb{R}^{q \times 1}$, $C \in \mathbb{R}^{1 \times n}$,
 $D \in \mathbb{R}^{m \times n}$, $e \in \mathbb{R}^{m \times 1}$ and $X \in \mathbb{R}^{n \times 1}$.

2.4 Hyperbolic Integer Programming Problem

Maximize $z = (CX + c_0)/(DX + d_0)$, $DX + d_0 \neq 0$
 subject to $AX \leq b$ $X \geq 0$ and $X \in \mathbb{Z}^{n \times 1}$, where $L \leq X \leq U$,
 $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^{m \times 1}$, $C, D \in \mathbb{R}^{1 \times n}$, $c_0, d_0, z \in \mathbb{R}$

3. Characteristics of the Random Search Methods

In the development of random search methods for mathematical programming problems, we have determined that the random search method requires the following, which we present here as the *characteristics of the random search approach*. The characteristics are as follows:

1. A clear definition of a successful solution in that situation is defined. This definition helps to establish bounds on the optimal solution.
2. An assurance factor for the final solution be pre-defined for that problem. This factor determines the number of trial runs needed to conclude the search.
3. At the end of every iteration when a successful solution has been obtained, the search region is reduced. For this purpose, it is essential to localize the search region for a successful solution. As the region of search reduces, the probability of finding an optimal solution increases.
4. An estimate of the total number of elements in the search region is made to make an initial estimate of the probability that a given solution is optimal, assuming, for optimality, points in the feasible region are uniformly distributed.
5. We are able to generate elements (solutions to the problem) in the search region by some random generator.
6. A criterion is established to terminate the iterative calculations.

4. Notations and Definitions

The following notation and definitions are required in the development of the random search method as described in section 5.

$A = [a_{ij}]$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$; $a_{ij} \in \mathbb{R} \forall (ij)$

$\hat{A}_j = \{\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jq(j)}\}$, $j = 1, 2, \dots, l$, where $\alpha_{jk} \geq 0 \forall (jk)$, $k = 1, 2, \dots, q(j)$.

Here $q(j)$ refers to the cardinality of \hat{A}_j for $j = 1, \dots, l$.

$F_i = \{X \in \mathbb{R}^{n \times 1}: X \text{ is a feasible point}\}$.

$H = \{x_j: x_j \in \hat{A}_j, j = 1, \dots, l\}$. H is an l -dimensional hyperbox with sides of length $q(j)$, $j = 1, \dots, l$.

ms = the number of search points required for the determination of a best integer solution for a given assurance $PROB$.

ne = An estimate of the number of search points in the search region H .

$PROB$ = The probability that at least one search point, in ms random searches, is an optimal solution ($0 < PROB < 1$).

X_0 = A given feasible value of X , $X_0 \in \mathbb{R}^{n \times 1}$, such that $x_j \in \mathring{A}_j, j = 1, \dots, l$, and $x_j \geq 0$ for $j = l+1, l+2, \dots, n$.

y_{jk} = a binary variable that can only assume a value zero or one $\forall(jk)$.

z_0 = the value of the objective function at a given feasible point.

Successful solution. In terms of optimum-seeking we would naturally regard the determination of the optimal solution as a success. In the context of the random search method we will modify the term *successful* to mean the determination of a feasible solution whose objective function value is at least equal to that we currently have, i.e. given our current *best* feasible solution X_0 with objective function value z_0 , another feasible solution X with objective function value z will be deemed successful if $z \geq z_0$. In general we may have many successful solutions $X_0, X_1, X_2, \dots, X_q$ such that $z_q \geq z_{q-1} \geq \dots \geq z_1 \geq z_0$.

5. A Non-Convex Optimization Problem

5.1 Statement of the Problem

Consider the following

$$\text{Maximize } z = CX \quad \text{subject to } AX \leq b, X \geq 0.$$

If we define $C_1 = \{x_1, x_2, \dots, x_l\}^T, X_2 = \{x_{l+1}, \dots, x_n\}^T$ such that $X = [X_1|X_2]^T$, and similarly define C_1 and C_2 such that $C = [C_1|C_2]^T$, then $z = z_1 + z_2$ where

$$z_1 = C_1X_1 \text{ and } z_2 = C_2X_2, C \in \mathbb{R}^{1 \times n}, C_1 \in \mathbb{R}^{1 \times l}, C_2 \in \mathbb{R}^{1 \times (n-l)}.$$

Consider further that $x_j, j = 1, 2, \dots, l$, can only assume a value from the finite set \mathring{A}_j , as defined in section 4. We call this NC1.

Suppose our current best feasible solution is X_0 with objective value z_0 where $z_0 = z_{10} + z_{20}$ then another feasible solution X_1 whose objective value is z_1 and $z_1 = z_{11} + z_{21}$ is *successful* if $z_1 \geq z_0$ when $z_{11}z_{21} \geq z_{10}z_{20}$. We verify this in *Appendix A*.

We extend this further to define the q th *successful solution* X_q with objective z_q , where $z_q \geq z_{q-1} \geq \dots \geq z_0$, which holds if $z_{1q}z_{2q} \geq z_{1q-1}z_{2q-1}$ is satisfied $\forall q > 1$, where $z_q^* = z_{1q} + z_{2q}$.

These problems arise when some variables in the LP formulation can only assume a value from some given finite set of values. This restriction makes the problem non-convex. The need for such restrictions may arise in allocation problems which are restricted by factors such as government regulations, availability of items or limited transport facilities. These situations may also occur in certain facility location problems where location may be restricted to one of several available options. An example of such a problem is presented in Section 7.

As an illustration of the application of the random search method to this type of problem we present an alternative formulation of NC1 in Section 5.2. This provides us with a search tool for part of the solution.

5.2 Zero-one Formulation of NC1

In order to apply the random search method we need to estimate the probability that a random point is optimal, see characteristic 4 section 3. We can easily achieve this by rewriting the problem NC1 as given in this section. The problem NC1 is equivalent to a binary optimization problem with $y_{jk} = 1$ if x_j takes the value α_{jk} , and 0 otherwise.

Thus the zero-one problem will be

$$\text{Maximize } z = CX$$

subject to

$$\sum_{j=1}^l a_{ij} \left(\sum_{k=1}^{q(j)} \alpha_{jk} y_{jk} \right) + \sum_{r=l+1}^n a_{ir} x_r \leq b_i, \quad i = 1, 2, \dots, m, \quad (5.2.1)$$

$$\sum_{k=1}^{q(j)} y_{jk} = 1, \quad j = 1, 2, \dots, l \quad (5.2.2)$$

$$y_{jk} = 0 \text{ or } 1, \quad \forall (jk), \quad x_j \geq 0, \quad \forall j \quad (5.2.3)$$

We call this NC2.

If the $q(i)$ values are large, NC2 can become a large problem computationally for application of the zero-one integer method. Hence we propose the use of the random search under these conditions.

5.3 Random Search Method for NC2

According to characteristic 4, we require an estimate of the number of integer points in the search region H. The number of points in the hyperbox represented by the zero-one variables will be given by $\prod_{j=1}^l q(j)$, thus $ne = \prod_{j=1}^l q(j)$. These points are not necessarily in the feasible region.

If the set of feasible points (F_i) is empty, the problem has no solution. Otherwise an optimal solution will exist. However, the problem may have more than one optimal solution. For simplicity, we consider here only the case of a unique optimal solution. It is clear that the probability p (that the search point under investigation is an optimal solution) can be given by $p = 1/ne$.

Now consider characteristic 6 for terminating the search. For this, an estimate of the value of ms is required. Since p is the probability that a search point under investigation is optimal, the probability that the current point is other than the optimal integer point is given by $(1-p)$. Thus the probability of not obtaining an optimal solution in ms number of searches is given by $(1-p)^{ms}$. Therefore the number of searches ms is a function of p and $PROB$ (the assurance factor), as defined by characteristic 2, is

$$ms = \lceil \ln(1-PROB)/\ln(1-p) \rceil.$$

For characteristic 1, we need to define the *successful solution* as discussed in section 3. This solution is used to satisfy characteristic 3, the reduction in the search region.

A *successful solution* X^* with objective value z^* reduces the feasible region when the constraint $CX \geq z^*$ is added to NC2. We call this problem NC3.

So at the q th successful solution, the constraint

$$CX \geq z_q^* \tag{5.3.1}$$

is added to NC2. We note that when $q = 0$, NC3 is equivalent to NC2. Therefore we have the algorithm as given in Section 5.4.

5.4 Random Search Algorithm for the Non-Convex Problem

1. Input n = number of variables
 m = number of constraints
 l = number of restricted x -variables
 $PROB$ = probability of assurance
 A_j for $j = 1, 2, \dots, l$
 $q(j)$ for $j = 1, 2, \dots, l$
 Initial feasible solution X and hence z i.e. X_0 and z_0 .
2. Calculate the value of ne ,
 If $ne = 0$ then goto 12 else calculate $p = 1/ne$.
3. If $p = 1$ then $\{q(j) = 1, \forall j$
 assign this to (5.2.2) and (5.2.1)
 if (5.2.1) is feasible then [solve the
 reduced LP of NC2
 Print "Optimal Solution and goto 12]
 else Print "Feasible Solution Does Not Exist" and goto 12}.
4. Calculate $ms = \text{Int}[\ln(1-PROB)/\ln(1-p)]$.
5. Set $s = 1$.
6. Call the *Random subroutine* to obtain y_{jk} values.
7. Assign the values to (5.2.1)
 if (5.2.1) is feasible then solve the reduced LP of NC3-this yields a new value of z (z^*)
 else goto 9.
8. If the feasible solution is *successful*
 then Print "Successful Solution" X^*, z^*
 and [assign $X_0 = X^*, z_0 = z^*, s = s+1$,
 Reduce search space i.e. for $j = 1$ to l do $q(j) = q(j) - 1$,
 Add constraint $CX \geq z^*$,
 goto 2.]
 else goto 9.

9. Set $s = s+1$.
10. If $s \geq ms$ then goto step 11 else goto 6.
11. Print "Solution" X_0, z_0 which is optimal with an assurance of *PROB*.
12. Stop.

Random Subroutine:

- for $j = 1$ to l and $1 \leq k \leq q(j)$ do
 (randomly generate y_{jk} values such that conditions (5.2.2) and (5.2.3) are satisfied.)

5.5 A Numerical Example

We illustrate the algorithm of 5.4 using the following example:

$$\text{Maximize } z = x_1 + 2x_2 - 13x_3 + 2x_4 + 6x_5 + 9x_6 - 2x_7 \quad (5.3.1)$$

$$\text{subject to } -2x_1 + 6x_3 + 2x_5 + x_6 - 3x_7 \leq 20 \quad (5.3.2)$$

$$-4x_1 + x_2 + 7x_3 + x_5 - x_7 \leq 10 \quad (5.3.3)$$

$$-5x_3 + x_4 + 3x_6 + x_7 = 60 \quad (5.3.4)$$

$$x_j \geq 0, \forall j \quad (5.3.5)$$

where x_1 is restricted to a value from $\{1,2,3\}$,

x_2 to a value from $\{0,1\}$, x_3 to a value from $\{1,2\}$ (5.3.6)

and x_4 to a value from $\{3,4,5\}$.

Iteration 1:

1. $n = 7, m = 3, l = 4, PROB = 0.7,$
 $\dot{A}_1 = \{1,2,3\}, \dot{A}_2 = \{0,1\}, \dot{A}_3 = \{1,2\}, \dot{A}_4 = \{3,4,5\},$
 $q(1) = 3, q(2) = 2, q(3) = 2, q(4) = 3$
 $X_0 = \{1, 0, 1, 3, 69, 0, 62\}^T, z_0 = 284.$
2. $ne = 36$ hence $p = 1/36$.
3. $p \neq 1$.
4. $ms = \text{Int}[\text{Ln}(1-0.7)/\text{Ln}(1-1/36)] = 42$.
5. $s = 1$.
6. *Random Subroutine* yields $y_{12} = y_{21} = y_{31} = y_{41} = 1$.
7. $X^* = \{2, 0, 1, 3, 73, 0, 62\}^T, z^* = 309$.
8. $z^* > z_0$ hence *Successful Solution*, $X_0 = X^*, z_0 = z^*,$
 $q(1) = 2, q(2) = 1, q(3) = 1, q(4) = 2$
 Add constraint $x_1 + 2x_2 - 13x_3 + 2x_4 + 6x_5 + 9x_6 - 2x_7 > 309$ to NC2.

Iteration 2:

2. $ne = 4$ hence $p = 1/4$.
3. $p \neq 1$.
4. $ms = \text{Int}[\text{Ln}(1-0.7)/\text{Ln}(1-1/4)] = 4$.
5. $s = 1$.
6. *Random Subroutine* yields $y_{11} = y_{21} = y_{31} = y_{41} = 1$.

7. (5.2.1) is not satisfied therefore infeasible.

9. $s = 2$.

10. $s < ms$ therefore goto step 6.

Iteration 3: yields $z^* = 330$ therefore *Successful Solution*, $X^* = \{3, 1, 1, 3, 76, 0, 62\}^T$.
 $ne = 0$ therefore the problem stops.

Therefore the *Optimal Solution with Assurance of 0.7* is (X^*, z^*) from iteration 3. (The optimal solution by Branch and Bound method is $z = 334$ with $X = \{3, 0, 1, 3, 77, 0, 62\}^T$).

5.6 Performance of Random Search Algorithm for Different Values of *Prob*

The algorithm for the nonconvex optimization problem has been applied to the example of 5.5 for different values of *PROB*. The algorithm was run 10 times for each value of *PROB*. The initial values of X_0 and z_0 used in all iterations are the same as that given in section 5.5. A summary of these results, including the mean values and ranges of z^* obtained are given in Table A.

In the cases where no successful solution has been found, the default value of z^* is 284. This occurred for the lower values of *PROB*. As would be expected the larger the value of *PROB*, the better the value of z^* obtained.

<i>PROB</i>	Max Number of Iterations (<i>ms</i>) required	Mean z^*	Range of z^* ([min,max])
0.1	3	309	[284, 334]
0.2	7	312	[284, 334]
0.3	8	324	[305, 334]
0.4	10	320	[299, 334]
0.5	14	324	[299, 334]
0.6	15	329	[299, 334]
0.7	28	331	[328, 334]
0.8	32	330	[301, 334]
0.9	50	334	[330, 334]

Table A - Solutions for the nonconvex optimization problem of 5.5 for different value of *PROB* (summary of 10 repetitions of algorithm of 5.4)

6. Ranking the LP Bases

In many real-world situations, where an LP formulation is valid, we may be interested not only in the best (optimal) solution but also the second-best, third-best,..., K-th best solution (in general $K > 1$). In such circumstances we may wish to determine whether there is much difference between the K solutions and hence decide whether to implement the best solution or one (or more) of the (K-1) solutions. Therefore the problem we consider is one of ranking the extreme points of the LP.

Consider the following problem of finding K ($K > 1$) bases of the LP {Max CX subject to $AX \leq b, X \geq 0$ } such that

$$CX_1 \geq CX_2 \geq \dots \geq CX_K,$$

where X_j represents the basic solution at the j th rank, $j = 1, \dots, K$.

It is obvious that CX_1 is the optimal solution for the given LP. Therefore our aim here is to find, not only X_1 but the bases X_2, X_3, \dots, X_K . We re-formulate this problem as follows:

Find X_1 using a linear programming solution procedure, then

for $l = 1, \dots, (K-1)$ do

$$\text{Max } CX$$

subject to $CX \leq CX_1 - \epsilon$, ϵ is a given small positive quantity

where X is an extreme point of $\{AX \leq b, X \geq 0\}$. (6.1)

For each l , this problem (6.1) is an extreme point mathematical programming problem, see Section 2.3. The solution of the problem can be approximated using the random search methods developed by Huynh and Kumar [11].

6.1 A Numerical Example

We reconsider the problem defined by (5.3.1)-(5.3.6), where the equality of (5.3.4) is replaced by the \leq inequality. We replace (5.3.6) by the following

$$x_1 \leq 3, \quad x_2 \leq 1, \quad x_3 \leq 2, \quad x_4 \leq 5$$

The random search method of [24] requires that the feasible region be bounded and non-empty. We therefore add the condition $z_0 \geq 0$ to this problem. The ranking of the basic solutions is given in Table B.

$(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$	z-value	Rank
(3, 0, 0, 0, 82, 0, 60)	375	1
(3, 1, 0, 0, 81, 0, 60)	371	2
(3, 0, 0, 0, 3, 20, 0)	201	3
⋮	⋮	⋮
⋮	⋮	⋮
⋮	⋮	⋮
(0, 1, 0, 0, 0, 0, 0)	2	18
(0, 0, 0, 0, 0, 0, 0)	0	19

Table B - The ranking of the feasible bases

6.2 Performance of Random Search Method for Different Values of *PROB*

Ten iterations of each value of *PROB* were performed on the above problem, and the feasible extreme points were determined and ranked according to the above process.

Table C includes a summary of the range of the *mean* of the number of feasible extreme points determined.

As would be expected, the range of extreme values found is largest for the smaller values of *PROB*.

Figure 1 is a plot of the value of *PROB* versus the mean number of extreme points found by the process. It tends to confirm the belief that the larger the value of *PROB* used the larger will the number of extreme points determined and ranked by the random search method.

<i>PROB</i> Value	Range of extreme points [Min, Max]	Mean number of extreme points
0.1	[3, 27]	16
0.2	[2, 23]	14
0.3	[2, 28]	12
0.4	[2, 18]	12
0.5	[7, 26]	18
0.6	[7, 26]	11
0.7	[8, 17]	13
0.8	[13, 24]	18
0.9	[18, 22]	19
0.95	[24, 25]	24
0.99	[27, 28]	28
0.9999	[30, 39]	33

Table C - Range and mean of the number of feasible extreme points found using the random search method

Mean N^o Extreme Points

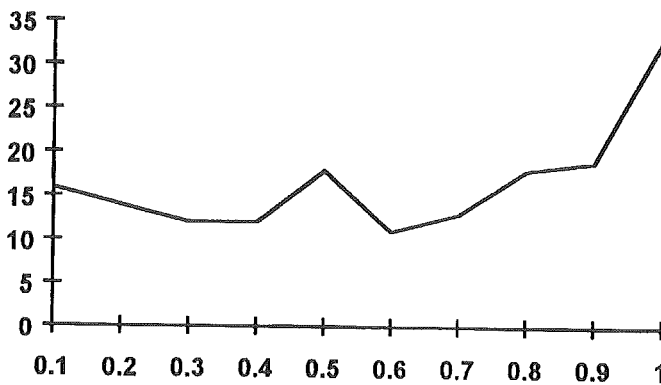


Figure 1 - Value of *PROB* (probability of assurance)

7. Applications of the Random Search Method

The random search approach may also be appropriate for the following situations

- (i) *A Location Problem* where an analytical solution tends to be difficult to determine, see Kumar [26].
- (ii) *A Single Machine Scheduling Problem*, see French [14].

7.1 A Location Problem

Australia Post recently installed Computer Services Terminals in post offices throughout the country. The number and location of these offices and their location is known. Each post office, depending upon the work it generates, may have one or more Counter Terminals (CT). Based on the past data, the client (Australia Post) is happy to assume that the number of CT's required in each office is known. Several of these CT's are assigned to a Concentrator, which has a connectivity capacity of handling 24 connections. A connection may handle one CT or several CT's through a device known as a jump device. Further, a single or group connection can be made in more than one way. The cost of connection and maintenance differ for each possibility. In all there are five mutually exclusive ways of connection.

The problem is determine the optimal location of the Concentrator sites where the cost of maintenance is a function of Concentrator sites.

Further we need to optimally distribute a given number of CT's in a given configuration to a given set of Concentrator locations. This is a function of the mode of connections. This is a large problem which can be seen in two phases:

- i. optimal distribution of CT's to Concentrators sites, and
- ii. optimal selection of the mode of connection.

This is an example of the *Star-Star Concentrator Location Problem (SSCLP)*. There have been a number of analytical and heuristic methods that have been developed for small to medium modules. A survey of these and related methods is available in Sridharan [27].

However, if the problem is large, an optimal or near-optimal solution is difficult to obtain.

Therefore in such a real-life situation an approximate solution with a guaranteed finite running time such as the random search method solution could prove useful.

7.2 A Single Machine Scheduling Problem

We now consider an n-job one machine scheduling problem, where the objective is to find the sequence which minimizes the maximal lateness (L_{max}). The main assumptions made in this problem are the same as those given in French [14], namely:

- (i) the number of jobs n is known and fixed;
- (ii) the number of machines m is known and fixed;
- (iii) the ready-to-start times r_j are known and fixed;
- (iv) the processing times t_j for jobs $j = 1, \dots, n$ are known and fixed;
- (v) no machine can process more than one operation at a time;
- (vi) no job can be processed by more than one machine at one time;
- (vii) no job preemption is allowed (i.e. jobs must be processed without interruption).

As started in Lawler [16] and proved in French [14], this $(1|r_j|L_{max})$ problems is NP-hard, except in the special case of unit-time jobs which is classified as easy [17]. A number of approximation techniques for this problem were implemented by Kise et al. [14].

In many real-world situations precedence relations exist among the jobs. This further complicates the model. Blazewicz et al [25] proposed a mathematical programming formulation of this problem. The random search method could be implemented here in order to obtain solutions with a reasonable assurance of being optimal or, at least, near-optimal. The random search method has the advantage over other techniques in the relatively small number of steps required to obtain a solution. This formulation is given in section 7.2.1. We illustrate using a small example in section 7.2.2.

7.2.1 Mathematical Programming Formulation of $(1|r_j|L_{max})$ with Precedence Constraints

Let x_{jk} be a binary decision variable which takes the value 1 if job j is the k -th job processed and 0 otherwise, and d_j be the due date of job j for $j = 1, \dots, n$. We also define r and s for jobs for which precedence conditions exist.

$$\text{Minimize } Y = L_{max} \tag{7.2.1}$$

subject to

$$\sum_{j=1}^n x_{jk} = 1 \quad (k = 1, \dots, n) \tag{7.2.2}$$

$$\sum_{k=1}^n x_{jk} = 1 \quad (j = 1, \dots, n) \tag{7.2.3}$$

$$\sum_{i=1}^k \sum_{j=1}^n t_j x_{ji} - \sum_{j=1}^n d_j x_{jk} \leq Y \quad (k = 1, \dots, n) \tag{7.2.4}$$

$$\sum_{i=1}^k \sum_{j=1}^n t_j x_{ji} - \sum_{j=1}^n (r_j + t_j) x_{jk} \geq 0 \quad (k = 1, \dots, n) \tag{7.2.5}$$

$$\sum_{i=1}^{k-1} x_{ri} \geq \sum_{i=1}^k x_{si} \quad (k = 1, \dots, n) \forall r, s \tag{7.2.6}$$

$$x_{jk} = 0 \text{ or } 1 \quad \forall (jk) \tag{7.2.7}$$

We note that constraints (7.2.2) and (7.2.3) are included so that assumptions (v) and (vi) of section 7.2 are satisfied. Constraints (7.2.4) and (7.2.5) refer to the lateness and non-preemption conditions of the model. Constraint (7.2.6) is added to allow for the precedence conditions of the model i.e. job r must be completed before job s .

7.2.1 Example

We consider the following single machine example, where the objective is to find L_{max}

(Job)	j	1	2	3	4
(Processing time)	t_j	5	7	6	4
(Due date)	d_j	6	8	9	10

subject to the precedence constraints that job 2 must precede job 4, and job 3 must precede job 1.

Case 1:

We assume $r_j = 0, \forall j$.

Using standard optimizer (EXCEL Solver):

Solution, using the formulation above, gives the order of jobs as 3 - 1 - 2 - 4, $L_{max} = 12$.

Using random search method:

PROB	L_{max}	Schedule
0.5	12	2-3-1-4
0.7	12	3-1-2-4
0.9	12	3-2-1-4

Case 2: Changing values of r_j :

(Both EXCEL and random search method with $PROB = 0.7$ give the same results)

From $r_1 = r_2 = r_3 = r_4 = 0$:

r_1	r_2	r_3	r_4	L_{max}	Schedule
1	2	3	4	22	3-1-2-4
4	3	2	1	22	3-1-2-4
4	0	4	0	20	3-1-2-4
10	10	10	0	42	3-1-2-4
20	0	20	2	52	3-1-2-4
100	0	100	0	212	3-1-2-4

(r_1, r_2, r_3, r_4) :

From	To	L_{max}	Schedule
(1, 2, 3, 4)	(4, 3, 2, 1)	22	3-1-2-4
(1, 2, 3, 4)	(10, 0, 10, 0)	32	3-1-2-4
(1, 2, 3, 4)	(100, 0, 100, 0)	212	3-1-2-4
(10, 0, 10, 0)	(4, 3, 2, 1)	22	3-1-2-4
(100, 0, 100, 0)	(4, 3, 2, 1)	22	3-1-2-4

Case 3:

Changing the Precedence: (starting with solution in case 1)

(i) Job 2 before job 3:

EXCEL Solver

$L_{max} = 12$, new schedule is 2 - 3 - 1 - 4

Random search

<i>PROB</i>	L_{max}	<i>Schedule</i>
0.5	12	2-3-1-4
0.7	12	2-3-1-4
0.9	12	2-3-1-4

(ii) Job 4 before job 3:

EXCEL Solver

$L_{max} = 16$, new schedule is 2 - 4 - 3 - 1

Random search

<i>PROB</i>	L_{max}	<i>Schedule</i>
0.5	16	2-4-3-1 (unstable here, failed to find a successful solution in 4 out of 10 trials)
0.7	16	2-4-3-1 (failed to find a successful solution in 1 out of 10 trials)
0.9	16	2-4-3-1

For a variety of small modules random search method obtains solutions comparable to a standard solver such as EXCEL. Naturally the accuracy and stability of the random method are better for the larger values of *PROB*.

The running time for the random search is faster than that of EXCEL for some instances (30% of trials) of *PROB* = 0.7 and most of those (68% of trials) for *PROB* = 0.5 Further work is needed on larger modules to determine the efficiency of the random search method for such single machine problems.

Appendix A

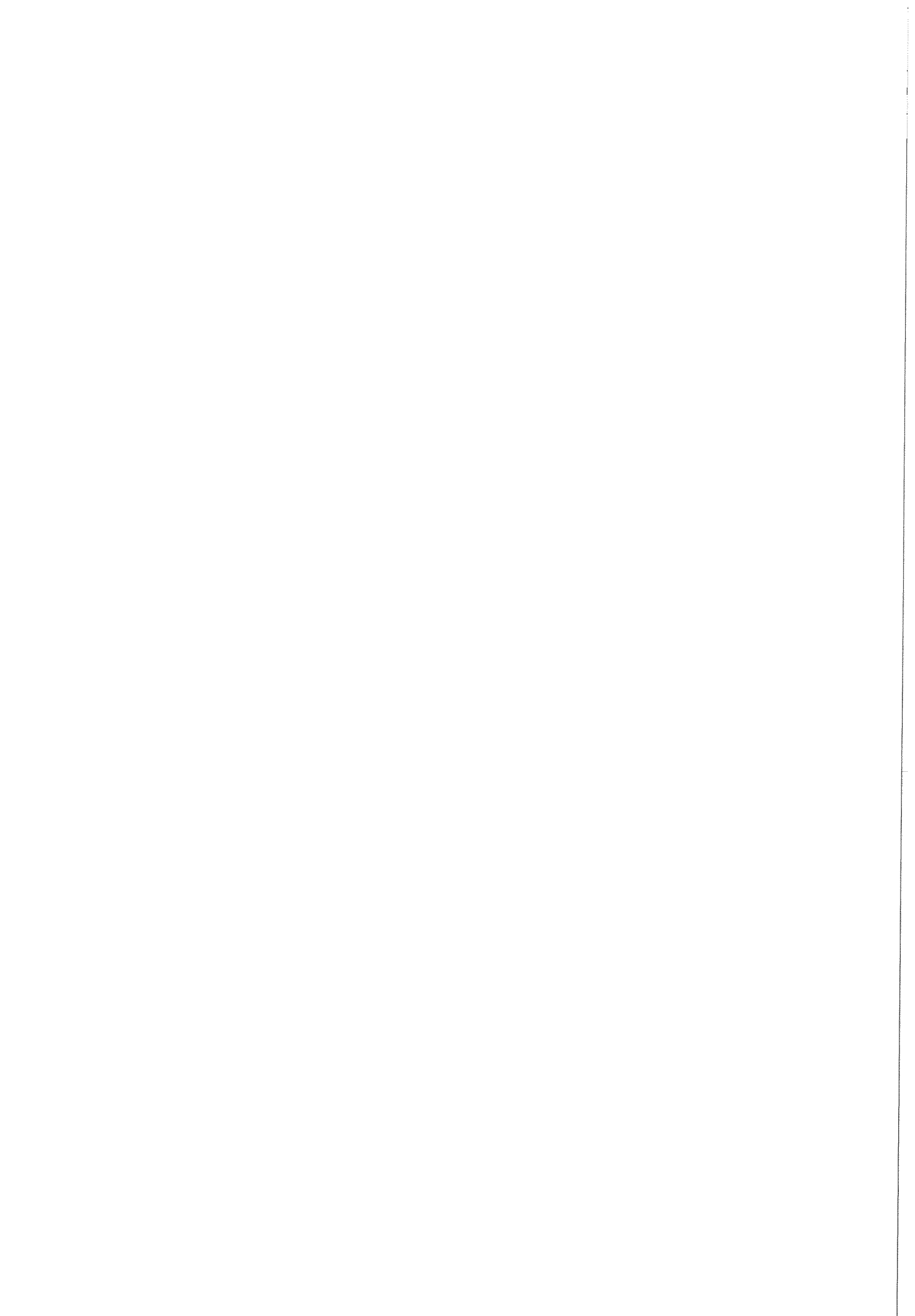
From $z_{11}z_{21} \geq z_{10}z_{20}$ it follows that $(z_{11}+z_{21})^2 + (z_{10}-z_{20})^2 \geq (z_{11}-z_{21})^2 + (z_{10}+z_{20})^2$ which is equivalent to $(z_{11}+z_{21})^2 - (z_{11}-z_{21})^2 \geq (z_{10}+z_{20})^2 - (z_{10}-z_{20})^2$
 Thus $(z_{11}+z_{21}) \geq (z_{10}+z_{20})$, ie. $z_1 \geq z_0$. Hence the proof.

Acknowledgements

Authors are thankful to the referee for his constructive suggestions.

References

- [1] Anderson, R.L., Recent advances in finding best operating conditions, *Journal of American Mathematical Association* 48 (1953) 789-798.
- [2] Brooks, S.H., A discussion of random methods for seeking maxima, *Operations Research* 6 (1958) 244-251.
- [3] Matyas, J., Random optimization, *Automat Rem Control* 26 (1959) 224-251.
- [4] Chichinadze, V.K., Random search to determine the extremum of the function of several variables, *Engineering Cybernetics* 1 (1967) 115-123.
- [5] Kjellstrom, G., Network optimization by random variation of component values, *Ericsson Technics*, 25 (1969) 133-151.
- [6] Anderssen, R.S., Global optimization, in: *Optimization*, Anderssen, R.S. et al (editors), Queensland University Press (1972).
- [7] Jacoby, S.L.S., Kowalik, J.S. and Pizzo, J.T., *Iterative method for nonlinear optimization problem*, Prentice Hall (1972).
- [8] Gottfried, B.S. and Weisman, J., *Introduction to optimization theory*, Prentice Hall (1973).
- [9] Anderssen, R.S. and Bloomfield, P., Properties of the random search in global optimization, *Journal of Optimization Theory and Applications* 16 (1975) 383-398.
- [10] Gaviano, M., Some general results of random search algorithms in minimisation problems, in: *Towards Global Optimization*, Dixon, L.C.W. and Szego, G.P. (editors), North-Holland/American Elsevier (1975).
- [11] Price, W.L., A controlled random search for global optimization, in: *Towards Global Optimization* 2, Dixon, L.C.W. and Szego, G.P. (editors), North-Holland (1978).
- [12] Kise, H. and Mine, H., Performances analysis of six approximation algorithms for the one-machine maximum lateness scheduling problem with ready times, *Journal of Operations Research Society of Japan* 22 (1979).
- [13] Solis, F.J. and Wets, R.J.B., Minimization by random search techniques, *Mathematics of Operations Research* 6 (1981) 19-30.
- [14] French, *Sequencing and scheduling: an introduction to the mathematics of the job-shop*, Wiley (1982).
- [15] Dorea, C.C.Y., Expected number of steps of a random optimization method, *Journal of Optimization Theory and Applications* 39 (1983) 165-171.
- [16] Lawler, E.L., Recent results in the theory of machine scheduling, in A. Bachem, M. Grottschel, B. Kote, *Mathematical Programming, the state of the art*, Springer-Verlag, Berlin (1983) 202-234.
- [17] Hamacher, H.W. and Queyranne, M., K best solutions to combinatorial optimization problems, *Annals of Operations Research* 4 (1985/86) 123-143.
- [18] Huynh, H.N., Kumar, S. and Connell, H.J., Random search method for integer programming, in: *Proceedings of the International conference on optimization techniques and applications*, Teo, K.L. et al (editors), National University of Singapore, Singapore (1987) 601-610.
- [19] Huynh, H.N., Connell, H.J. and Kumar, S., Random search method for mixed integer programming, in: *Proceedings of the 8th National Conference of the Australian Society for Operations Research*, Kumar, S. (editor), Royal Melbourne Institute of Technology, Melbourne (1987) 109-118.
- [20] Raghavan, P. and Thompson, C.D., Randomized rounding: a technique for probably good algorithms and algorithmic proofs, *Combinatorica* 7 (1987) 265-274.
- [21] Huynh, H.N. and Kumar, S., Random search method for hyperbolic integer programming, in: *Proceedings of the 9th National Conference of the Australian Society for Operations Research*, Hoffman, D.L. (editors), Canberra (1988) 93-99.
- [22] Mohan, C. and Shankar, K., Reliability optimization of complex systems using random search technique, *Microelectronics and reliability* 28 (1988) 513-518.
- [23] Shankar, K., Khattri, K.N. and Mohan, C., Inversion of gravity data by random search technique, *Journal of Association of Exploration of Geophysics* 10 (1989) 153-170.
- [24] Huynh, H.N. and Kumar, S., A random search method for extreme point mathematical programming, *Asia-Pacific Journal of Operations Research* 7 (1990) 30-45.
- [25] Blazewicz, J., Dror, M. and Weglarz, J., *Mathematical programming for machine scheduling: a survey*, *European Journal of Operations Research* 55 (1991) 283-300.
- [26] Kumar, S., *Electronic communication optimization problem*, Technical Research Report prepared for Australian Postal Corporation (1991) 1-11.
- [27] Sridharan, R., The capacitated plant location problems: a survey, in: *Decisions Sciences: Tools for Today*, *Proceedings of the 12th National Conference of the Australian Society for Operations Research*, Sutton, D.J., Pearce, C.E.M. and Cousins, E.A. (editors), Adelaide (1993) 609-647.



STRATEGIC APPROACHES TO CLIMATE CHANGE AT COUNTRY LEVEL - FOCUSING ON GREENHOUSE GAS ABATEMENT

Pedro Borges

Universidade Católica Portuguesa
Rua Diogo Botelho, 1327
4150 Porto - Portugal

Lene Sorensen

Center for Tele-Information
Technical University of Denmark

Arturo Villavicencio

UNEP Collaborating Centre on Energy and Environment
Risø - Denmark

René Victor Valqui Vidal

Institute of Mathematical Modelling
Technical University of Denmark
2800 Lyngby - Denmark

Abstract

National climate change costing studies would generally benefit from the strengthening of their strategic dimension, especially in the case of developing countries. This paper identifies shortcomings of current approaches and persisting organisational and process problems. An institutional set-up is suggested, as well as an attitude towards strategy development and planning.

For the planning process two approaches are illustrated. Both are based on the same well known tool for strategy formation (SWOT), but use it in different ways. The first is a classical approach that focuses on the higher planning sphere, while the second is a mixed approach intended to seek agreement between decision makers and members of a particular working team on climate change studies. The proposed measures and tools allow for long term planning and provide a better framework for integrating national studies and decision making processes.

Resumo

Os estudos para estimação dos custos associados ao fenómeno de aquecimento global e à limitação de emissões de gases de efeito de estufa podem beneficiar de um reforço da sua componente estratégica, especialmente em países em vias de desenvolvimento. Este artigo identifica limitações e problemas organizacionais observados em alguns desses estudos.

Sugere-se a integração de métodos de estruturação de problemas nos processos de planeamento e criação de estratégias. Esta medida pode auxiliar o planeamento de longo prazo e criar um quadro de referencia para a integração dos estudos na estrutura dos órgãos de decisão dos respectivos países.

Keywords

Strategic planning, SWOT, mixed approach, abatement costing studies, mitigation strategies, developing countries, participation

1. Introduction

"[...] policies and measures to protect the climate system against human-induced change should be appropriate for the specific conditions of each party and should be integrated with national development programmes [...]".

(United Nations, 1992)

Climate Change has become an item on the agenda of most governments since the late eighties. By "Climate Change" it is meant change of climate attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is, in addition to natural climate variability, observed over comparable time periods [13]. Climate change is one of the most complex issues mankind currently faces, and is connected with many vital sectors of the world economy and many basic human activities. Climate change is mainly attributed to emissions originating from fossil fuels combustion, which are expected to increase globally even in the most optimistic scenarios. Therefore, the term used for addressing the problem is *climate change mitigation*, meaning that the future emissions and their impacts can be made less severe than the projected estimates, but not reduced from current levels.

The use of fossil fuels is expected to increase in developing countries, since energy use is associated with economic growth. However, those countries do not usually prioritise environmental concerns in their actions, since there are many other essential and urgent issues they must address. Many countries are striving for survival and it is not acceptable that their development is delayed by measures from the international community to mitigate climate change, since they were not the main contributors to the problem. About 80% of the earth's energy resources are used by only one fifth of the population [3]. On the other hand, disregarding the opportunity of taking environmental issues into account early in the development process might be an unsound strategy. In the long range, the poor will suffer most from the negative environmental impacts and pollution, since they do not have the resources to adapt to a changing environment. Energy and environmental policies and planning issues should be considered together with the national economic development objectives. The question is then, how to design a process that the stakeholders will accept and that might change their attitude towards including mitigation in their development plans. With such a process one can only hope that countries will follow sensible paths towards better situations.

This paper addresses the difficulties that arise around carrying out mitigation studies and implementing mitigation strategies. The focus is on institutional set-up and strategy development and planning when dealing with mitigation issues and the application of methodologies that can help structuring problems (referred to as problem structuring methods) in this context. This paper addresses no specific country, but is based on some experience of working with developing countries [12].

The next section (Section 2) introduces the problem of climate change, together with the international reaction it motivated. Section 3 gives an illustration of a methodological approach for applied abatement analysis used in several countries. Section 4 summarises circumstances

observed in developing countries that affect their response to climate change and to carry out particular studies. Section 5 discusses how the current situation could be improved and suggests a different institutional arrangement together with the possibility of using problem structuring tools in the strategic planning process. Two examples are given in Section 6, a more classical top-down hierarchical approach and a mixed one. Finally, Section 7 includes some conclusions and comments.

2. Background

2.1 Greenhouse effect and Climate Change

Today's knowledge of climate is neither enough to predict the local weather within more than a few days nor sufficient for long term regional and global forecasts. Climate is ruled by complex mechanisms and is influenced by many factors, like solar radiation, earth's orbit, effects of volcano eruptions, the composition of the different layers of the atmosphere, and the interactions with living organisms. Effects of change in these and other factors might accumulate and not show, or prevail, until many years after the originating event occurred. Climatic changes can also trigger other processes or feedback loops, about which much knowledge remains to be gathered and validated [7].

In this century, the man-induced changes in the atmosphere appear capable of irreversibly altering "natural" climatic patterns. The gross part of this change might occur by modification of earth's radiative balance through enhancement of the *greenhouse effect*. The mechanism of this effect consists in that the short-wave solar radiation that reaches earth's surface is partially reradiated back in longer wavelengths, which might be absorbed or reflected back to the surface by the greenhouse gases in the atmosphere, thus reheating the surface. The relevant greenhouse gases (GHG's) are: water vapour, carbon dioxide, methane, nitrous oxide, CFC's and tropospheric ozone. Aerosols also influence climate but are thought to have an overall cooling effect. The natural physical processes taking place on the surface and atmosphere are responsible for the earth being on average about 34 degrees centigrade warmer than if it was not protected by the atmospheric layer.

If earth's surface is warmed up by even a few degrees, weather patterns and extreme event distribution might be modified and the sea level will rise. The rise of the sea level will be caused both by faster melting in the poles and glaciers, and by thermal dilatation of the oceans. This rise is currently expected to be only of a few millimetres per decade, but that could be enough to create problems, particularly to low-lying coastal and island areas.

2.2 Carbon Dioxide

Records show that the atmospheric concentration of CO₂ increased by about one third since the beginning of the industrial revolution. Even if the impact of that increase is not accurately known, at least it is certain that the risk of it adding global warming to our planet exists. Carbon

dioxide is also the most important GHG, and is thought to be responsible for more than 50% of the total radiative forcing.

Beyond being the most important GHG, anthropogenic emissions of CO₂ are related with the use of fossil fuels. This made it possible, at least at first, to focus primarily on the energy sector. Still, this is not a simple task, since the burning of fossil fuels may take many different forms and is complicatedly associated with many human activities. The CO₂ emissions from use of fossil fuels are mainly due to the use of coal, natural gas and oil. Of these, natural gas gives the lowest emissions while coal has the highest emissions per unit of energy produced.

2.3 International Response

In 1988 the United Nations Environment Programme (UNEP) and the World Meteorological Organisation (WMO) established the Intergovernmental Panel on Climate Change (IPCC) which in 1990 presented its first assessment report describing the existing knowledge and scientific consensus on climate change issues.

The Rio conference on Environment and Development in 1992 was the stage for the signature of a UN Framework Convention on Climate Change (FCCC), by more than 150 countries. The objective of the FCCC is to achieve "stabilisation of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system. Such a level should be achieved within a time-frame sufficient to allow ecosystems to adapt naturally to climate change, to ensure that food production is not threatened and to enable economic development to proceed in a sustainable manner" [13]. The FCCC recognises that most of the historical and current global emissions have originated in developed countries. The convention also foresees mechanisms for transference of resources from industrialised parties to developing countries to help in reducing emissions, which demands for studies on the cost of abatement strategies and agreement by the parties on the used methodologies. The FCCC came into force in 1994. It requires parties to prepare national reports on greenhouse emissions and sinks, and serves as legal basis for the international development of responses to the climate change problem. More recently, the Kyoto protocol [14] set the first international abatement commitments under the FCCC.

Greenhouse gas emissions assume multiple forms and interact with the economy in complex ways. Because of that, they are understood differently by different groups and it is not bizarre that the first studies on cost estimation for limitation of GHG emissions have arrived to apparently contradicting results because of different approaches. To work around this difficulty, UNEP started a project on "National Greenhouse Gases Abatement Costing Studies" with the goal of producing a methodology for carrying out comparable and transparent costing studies on reducing GHG's. That methodology comprises three main elements:

- Definition of key concepts and terminology;
- A set of quantitative assumptions; and
- An analysis structure.

3. The UNEP guidelines for abatement costing studies

The UNEP costing project resulted from the need of comparability between country assessment studies on greenhouse abatement cost, and aims at developing a methodological framework for those studies through the support, cooperation and interaction with teams carrying out the national studies.

The proposed process for reaching the desired uniformity and explainability follows the steps of defining a reference scenario (baseline), identifying abatement options, constructing abatement scenarios, and assessing their impact. The structure of applied abatement analysis proposed by UNEP is presented in Figure 1.

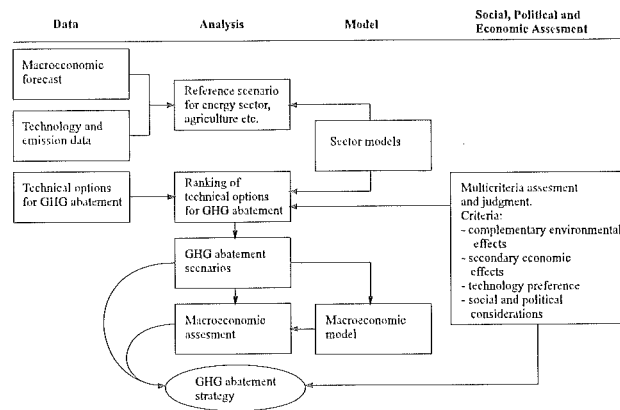


Figure 1 - UNEP's structure of applied abatement analysis procedure [12]

National studies intend to estimate GHG abatement costs for one or several scenarios following the general assumptions specified by the UNEP guidelines. However, they are specific to each particular country and costs must therefore be regarded from that country's point of view. For that reason the proposed framework emphasises uniformity in establishing scenarios and performing the analysis, rather than in data and modelling issues, which will vary according to the country.

The term *abatement cost* is defined as the long-term cost to society of implementing a GHG reduction strategy compared with a hypothetical reference case where no effort is made to reduce GHG emissions. Cost measurement includes investment costs, operation and maintenance costs, and fuel costs for the individual options. Particularly interesting are the so-called *no-regret* options, that is, options whose benefits exceed the required investments.

As seen in Figure 1, the UNEP approach recognises and explicitly states the importance of considering multiple criteria for ranking the technical options and evaluating abatement strategies. There is no doubt that country decision making processes will be conditioned by other factors than direct cost and greenhouse gas abatement potential. Criteria and constraints other than direct cost were only addressed in broad qualitative terms and generally not incorporated in the cost estimates. The UNEP guidelines suggest the analysis of "joint products and externalities" together with information on national priorities and constraints in order to reveal possible interactions between least-cost abatement options and national development plans.

Even if guidelines exist, it is not straightforward that they are accepted or taken into consideration in the national planning processes. Suppose a country that acknowledges the need for a strategy document on climate change mitigation. Two extreme cases of non collaboration between decision makers and the analysts making a country study could occur. In the first, a government official produces a document on mitigation strategies based solely on general information on country policy and without regard for constraints or technical aspects that must be present in national strategies. That is, a political document that serves no purpose in practice, other than being acknowledged as existent (decorative planning). In the second, a technical team is assigned the task of conducting a mitigation national study and receives external funding. The mitigation options included in this study reflect neither the priorities of the government nor the implementation difficulties. Therefore, the study does not rise interest from decision makers who see it as an academic exercise and do not share the assumptions of the people who made it (academic planning). If GHG mitigation strategies are to be meaningful the planning process must combine the decision makers' views and motivations with the essential rigour in technical evaluation of requisites, costs and impacts.

4. The circumstances of developing countries

4.1 Planning horizon

Climate change studies address long-term decisions and impacts, yet the economic and political uncertainties in the long term are so large, that national forecasts are generally precariously balanced and almost isolated guesses. Abatement studies are based on a single baseline scenario and the UNEP studies typify different ways of building it, but no sensitivity analysis is attempted on the impact of deviations and variations in the forecasted/projected trends. How would the most probable scenario change if the projected evolution of a few variables was slightly different? And more important, would the interpretation and conclusions of the country study be affected?

National development scenarios tend to merely quantify alternative outcomes of obvious, simplistic and conflicting situations, without indications upon which decision-makers could understand the nature of uncertainties and how to come to grips with them. In general, "high", "middle" and "low" scenarios are in fact straight-line extrapolations, only with different slopes. They do not operate within the constraints of what is politically feasible, economically reasonable and technically attainable.

Although mitigation analysis must be embodied into the national strategies, governments in many developing countries have almost abandoned the practice of long term planning. This fact, apparently absurd, is explained by two main reasons. One is that the sense of emergency caused by a variety of almost permanent crisis (inflation, foreign debt, rising poverty, urban migration, social unrest, political instability, etc.) draws governments into day-to-day management of the crises, putting aside any preoccupation on longer term issues. The second

reason is the idea that market liberalisation policies will automatically create the conditions for an efficient allocation of resources and thus turn long-term planning useless.

In many countries, economic and planning teams have been or are being dismantled as part of structural adjustment programmes. In South America, for example, national planning teams have been dismantled in countries like Brazil, Argentina, Peru and others. This so-called incrementalist approach resulted that, in the best cases, only short term (2-5 years) projections could be available for some macroeconomic aggregates (GDP growth rate, foreign debt service, trade balance).

4.2 Institutional Setting

Environmental issues suffer often from fragmentation of responsibilities that are spread over several government agencies and ministries. The establishment of special commissions (commissions for environment, for climate change, for global change, etc.) does not guarantee a successful coordination of policies and strategies. The existence of a complex system of rules and regulations by which competencies, tasks and responsibilities are divided among ministries, directorates and departments, complicates the coordination tasks. Experience shows that projects that involve technical and financial assistance are likely to antagonise the participating agencies rather than enhancing cooperation.

In several countries, national studies entailing the involvement of a network of agencies are to be coordinated by meteorological institutes or by ad-hoc commissions. Often, those institutions lack the experience and the expertise for conducting or even coordinating a nationwide study. Furthermore, it is utopian to expect that weak institutions, lacking the mandate and the technical capacity, will be able to contest ongoing development strategies and advocate for more sustainable development policies.

4.3 Objectives and mitigation option choices

What are the objectives of the country teams when designing mitigation strategies and assessing abatement scenarios? Different institutions will suggest different strategies because of their different motivations. What are usually regarded as national objectives, i.e., the increase in economic growth, the promotion of rural development, the increase in self reliance, improved social and physical infrastructures, etc., represent the logical aspirations of developing countries towards their ideals and remain the same independently of possible development patterns or scenarios. The interests and opinions on how to reach those consensual general national objectives differ, thus affecting their choices.

Different views mean different assumptions and apriori choices that are generally not presented together with the mitigation scenarios. Mitigation studies undertaken in countries with comparable development levels and similar characteristics show a great divergence in the range of options analysed. It seems that the selection of mitigation technologies are mainly dictated by

the analyst's imagination knowing that she/he will never have to take the decisions necessary to implement those options.

If the political priorities of governments and governmental agencies are not considered when selecting abatement options, the proposals of abatement analysis will conflict for resources with other measures. Analogously to the manner the UNEP guidelines assess no-regret options based on their direct cost [12], one should also prioritise abatement options that bring benefits in other public planning priorities and so would be more easily implemented. This is sometimes called double-dividend.

5. A broader approach to GHG studies at national scope

This section presents suggestions to address some of the issues that limit the quality and impact of climate change studies. The essential idea is to try to narrow the predominant gaps between national climate change studies, decision making and implementation.

Three purposes for this approach should be underlined:

- Create the conditions for managing long term scenarios and plans;
- Bring the priorities of decision makers into mitigation studies; and
- Widen the distribution and discussion of national studies conclusions so as to increase and support environmental concern in national decision making.

5.1 The national environment and the available capacities

To avoid the risk that mitigation studies become limited to isolated exercises and remain confined within a very narrow sphere of influence, any approach should attack the problem (specific for each country) of identifying, promoting or designing the "policy carrier" for climate change issues. This policy-carrier is an appropriate institution (or institutional set-up) which

- has legal mandate and interest on climate change issues and follows the FCCC,
- is able to generate and process policy information, and
- has the political power to influence decisions and materialise policies.

In most countries, no such "policy carrier" can be identified. For this reason one should look for a process to fulfil some of the existing needs and assume at least partially the role of the "policy carrier".

The first task is to identify partners in the country that: have interest in the mitigation issues; would be responsible for the implementation of mitigation options; possess the technical expertise; or have legal competence in the area. The planning horizon and time scope of mitigation studies are long and may run even while political changes occur. Therefore, it is important to identify partners accepted as credible and independent by a broad national audience. Some partners must be government representatives, able to convey the relevant information to the political scene and privileged contacts for adjusting lines of work to different eras and national priorities. Other partners might be non-governmental organisations,

representatives of GHG emitting sectors, environmental organisations, social partners and the scientific community. The ideal institution would be able to receiving a broad national confidence vote in climate change issues and have reasonable independence from the political power itself. All these partners should have the possibility of contributing and receive information on the national climate change activities.

From this group of partners a group of representatives should be selected, if possible combining cross-ministerial participation and independent expertise. It is clear that this is not feasible in many countries for some of the reasons mentioned in Section 4. This could be seen as an ideal situation, countries should get closer to.

Let us assume that one gathers this group of representatives in a sort of *Climate Change Commission* that represents the interest groups on climate change issues and coordinates actions to pursue the following:

- Maintain a plan related to climate change;
- Accompany the international discussions and agreements on climate change;
- Include national priorities and constraints into climate change studies; and
- Bring the concern on mitigation issues closer to the decision arena.

Note that not all the members with seat in this Climate Change Commission (CCC for short) have the same motivation or interest in taking active part in the technical details of particular national studies. Nevertheless, they might aspire to influence and contribute or have interest in being fed with reliable information regarding mitigation. Because of the different purposes, expectations and ideas, it is not expected that they always agree on the involved topics. A good deal of conflict will exist.

Why should a new entity be inserted between government and climate change studies, if the objective is to bring decisions closer to the study purposes and results? The CCC should do a job that currently is not satisfactorily done and is dispersed. It should also possess as much scientific competence as possible and follow long term issues and objectives that are not pursued by day-to-day or mandate-to-mandate governmental practice.

The particular operation of the CCC, or equivalent institutional arrangements, will be different from country to country according to the specific conditions. The interaction of the CCC with the political decision system is not to be neglected, since collaboration and acceptance from the government are essential for its purposes. For the sake of concreteness, the remainder of the paper refers to a hypothetical CCC.

Figure 2 displays a possible design for the relation between the CCC and a team elaborating a specific national study on mitigation issues. The study team receives guidelines from the CCC (in the form of suggestions, questions, criteria, issues to investigate, etc.) and is accompanied by a group that allows feedback and might include external consultants (as often happens, since most country studies are part of bilateral aid programmes.)

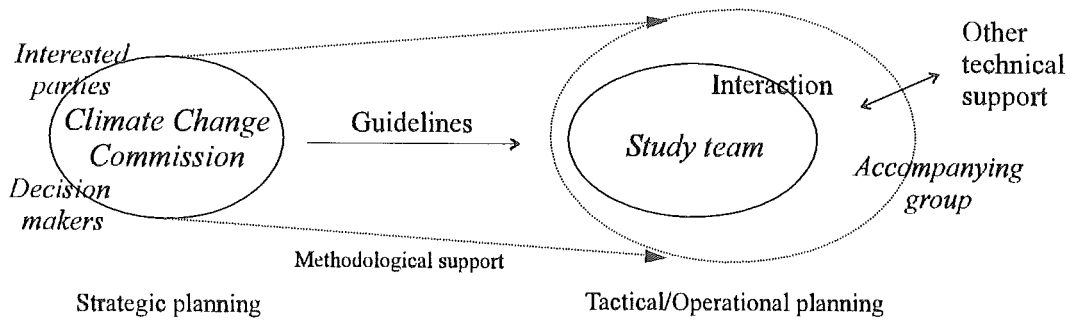


Figure 2 - Relation between CCC and the working team for a particular study

The question remains on how to achieve a process that might reflect national objectives in country studies and to bring mitigation analysis closer to the decision stage. The existence of communication between decision makers and study teams is essential and some sort of CCC is necessary for this purpose but does not guarantee success. Still to be dealt with are the questions on how to cope with multiple criteria, uncertainty, divergent points of view, and how to develop strategies.

5.2 Roles for planning

Following Mintzberg [6], *planning* is defined as a formalised procedure to produce an articulated result, in the form of an integrated system of decisions. Formalisation is the central issue and usually implies to structure, to articulate, and to "rationalise" the process by which decisions are made and integrated in organisations. Articulation means that the product of planning, the plan, must be clearly and explicitly labelled by words and numbers. For an organisation, the concept of strategy is related to its environment and can be synonymous of a position, a plan, a ploy, a pattern or a perspective.

An important aspect is that planning activities are social processes involving decision makers, planners, methods, and problems [15]. In this process there are risks for contradictions at different levels. A basic condition for the success of a planning process is that the participants are able to understand it, which should constrain the levels of abstraction allowed in the different analysis stages. This must not be overlooked in the case of climate change, where individuals with significantly different roles and background should interact.

Following Emery [2] we suggest that the planning process for country mitigation strategies be formally divided into two main stages:

- Strategic stage
- Tactical and operational stage.

It is assumed that the CCC retains a strategic perspective, while particular teams (eventually within the CCC) have operational tasks (Figure 2). Planning is not an act or a set of isolated acts, but a process. The strategic stage has a continuous role over time, needing updating and maintenance to accompany national and international evolutions in variables and factors related to the climate change issue.

The objectives of the strategic stage could be:

Situational Diagnostic

- Keep an inventory of available expertise on climate change related issues.
- Accompany international climate change efforts.
- Keep and update a national inventory on GHG emissions and sinks.

Development of strategies

- Define relevant guidelines containing information on relevant public priorities and objectives, that should be taken into analysis in the studies.
- Agree on a minimum set of practical issues for carrying out the operational tasks:
 - Decide aims and methodological framework of particular studies;
 - Determine who will coordinate and compose the working team;
 - Find and allocate resources; and
 - Participate in the definition of a time plan.
- Decide on the aspects of announcement, participation, dissemination and discussion of the studies.
- Design implementation and follow-up systems.

The tactical/operational stage focuses on more specific aspects:

- Conduction of the technical parts of studies; performing detailed analyses
- Interaction between the CCC and the study teams, maybe in the form of project status information and comments.
- Dissemination of information.

The remainder of this paper addresses mostly the strategic planning issues.

5.3 Participation, information

To avoid that mitigation studies remain isolated modelling exercises, the publication and dissemination of country study results must be carefully planned. Supposing that a Climate Change Commission exists and that a working team is designated within it to carry out the study, the following participation and communication mechanisms should exist:

- Interaction/influence between working team and the CCC,
- Interaction/influence between CCC (working team) and the decision makers.

The manner of doing this depends certainly on the particular organisation and culture of the particular country. Three functions are important, here referred to as:

- Consensus workshop;
- Midterm interaction; and
- Consensus conference.

The consensus workshop is the inaugural sharing opportunity between the CCC and the team elaborating the study. This workshop includes phases of self presentation and information about specific aspects and objectives of the study, so that the points of view underlying the definition of criteria, options, and strategies are understood and accepted by the

participants. This is a requisite for them to be able to argue and discuss on the same basis. Other phases could include partial meetings, discussion panels, and questions and answers panels. The result of this workshop should be a document containing ideas and guidelines for the study, accessible to the public. This first document should include the agreed mechanisms for midterm interaction and comment processing.

The consensus conference would be held at the end of a particular study, allowing for questions by members of the CCC and potentially any interested citizens (including the media) to the study team. It could also result in a document written by the CCC, and a summary written in a language that allows for distribution to the general public. Appropriate funding for this conference should be inscribed in the budget of the project from its start.

These mechanisms are to be understood only as possibilities that need to be reshaped to adapt to the realities of particular countries. They are too ambitious for most developing countries, where institutions do not function and the concepts of open discussion and democracy are not adopted. The guiding direction should however be increasing dialogue and conveyance of ideas to the political spheres, which would invariably be beneficial. The use of tools allowing for participation, such as problem structuring methods mentioned ahead, might help steering towards that direction.

The possibility of public debate or reaction is important and too often decisions that affect entire populations are made by a few and without adequate knowledge. This risk can be illustrated with an example: There are many models that attempt to model climate change and mitigation issues or their impacts. There is a risk that those models are regarded by developing countries as opaque tools they can get for free (e.g., getting bilateral help from a developed country conditioned to the use of a particular model) and used without understanding their assumptions or being given the underlying mechanisms. This results in a validation deficit that might have perverse consequences and does not really bring know-how to the developing country, but accentuated dependency.

A programme of continuous monitoring, study update, and follow-up must also exist under the climate change commission.

5.4 Problem structuring methods for planning

During the last decades many methods have been suggested that try to give structure and help modelling complex situations. Some examples are [9]: cognitive mapping, soft systems methodology, strategic choice, robustness analysis, metagame analysis, and hypergame analysis. They have different applicability and characteristics, even if they share the attribute of trying to structure unstructured information into some kind of model [10]. This type of approaches could be helpful in designing strategies for climate change related issues, as well as for identifying relevant criteria to be considered.

For the sake of illustration the idea of the TOWS matrix, proposed by Wehrich [17], is employed in this paper. This tool gives the ownership of the problem to the actors, still

allowing them to manage the information they disclose (in particular, for political partners). Its original name is TOWS [17], but the designation SWOT, standing for Strengths Weaknesses Opportunities and Threats is more common.

The SWOT matrix is a structuring tool for situational studies that systematically analyses the external and internal environments of an organisation. In our example the organisation is a whole country. It involves the assessment and appreciation of the external factors and from those identifies opportunities and threats posed to the organisation by the external environment. Likewise, the internal factors are used to list strengths and weaknesses inherent to the current status of the organisation. Representing strengths, weaknesses, opportunities, and threats in a tabular format gives origin to the SWOT matrix, represented schematically in Figure 3.

The SWOT matrix suggests four different ways for generating strategies by combining the minimisation of threats, minimisation of weaknesses, maximisation of strengths, or maximisation of opportunities. The different possibilities are explained next from a general point of view (see Figure 3) and used later in examples.

The maxi-maxi strategies attempt to maximise the opportunities by making use of and improving the internal strengths, while mini-maxi strategies try to do the same by minimising the internal weaknesses, eventually transforming weaknesses into strengths. The maxi-mini strategy intends to minimise the impact of the threats by improving the internal strengths while mini-mini strategies also intend to minimise threats but do it by minimising weaknesses.

	Strengths	Weaknesses
Opportunities	Maxi-maxi strategy	Mini-maxi strategy
Threats	Maxi-mini strategy	Mini-mini strategy

Figure 3 - The SWOT matrix and possible strategy classes

This description of the SWOT matrix is enough for the purpose of this paper but is too short as an introduction to the method. It serves, however, to illustrate the conceptual simplicity behind this tool which allows easy familiarisation and sense of control over the method (see further [8]).

The SWOT approach is commonly used in many organisations. It favours dialogue in the planning process by including a model that is shared by the participants.

6. Approaches to a strategic planning process

To avoid the risks of being academic or decorative, the planning process must combine the decision makers' views and motivations with the essential rigour in technical evaluation of requisites costs and impacts. The next section describes two possible approaches to the strategy

development of mitigation policy. Both are illustrated with an example. The first example only presents issues to be considered, while the second includes detail for a particular country.

6.1 A Classical Approach (Top Down)

Here, the planning process addresses first the highest level, closer to the decision makers, and only in a later stage deals with specific and more technical issues.

This type of approach is mainly used in hierarchical organisations. In this kind of process the task of planning is relatively easy, while the implementation may be complicated or impossible.

6.1.1 Example of a classical approach using SWOT analysis

If the climate change commission of a developing country wants to propose general strategies for addressing mitigation, several internal and external factors can be identified and classified according to the SWOT matrix.

In the following, topics are presented that could be part of the analysis of a country:

External opportunities

Some opportunities derive from the Framework Convention on Climate Change itself. The FCCC acknowledges that developed country parties shall provide new and additional financial resources to meet the agreed full costs incurred by developing country parties in complying with their obligations under the convention. It creates a mechanism (GEF) for the provision of financial resources on a grant or concessional basis, including for the transfer of technology.

Availability of external funding for carrying out national studies exists, plus the availability of multi-lateral funds for covering the incremental costs resulting from the implementation of climate change mitigation options. Additional resources are also available from bilateral assistance.

Joint implementation actions [5] are an opportunity for obtaining additional resources for priority projects and gain access to new technologies. The FCCC recognises that abatement options should be ranked taking cost effectiveness into account. The abatement cost per ton of carbon dioxide is higher in developed countries than in developing ones, which makes it more cost effective to use resources in developing countries. Through *joint implementation actions* the developed partner supports mitigation actions in a developing country and both partners would gain *abatement credits*.

Some actions could bring social and economic benefits through improving infrastructures, fomenting employment, or other environmental improvements. Opportunities may be identified for accelerating the transfer of new technologies. This could improve the conditions for technological development within the countries.

External threats

Developing countries perceive climate change concerns as a risk for reallocating external financing resources to GHG related projects. Since most of the development agencies and

donor countries are undertaking cuts or at least freezing their budgets for development assistance, there is a fear that traditional areas which benefited in the past from external funds will be affected by the diversion of scarce resources to climate change activities. This is obviously seen as a threat by developing countries, where mitigation plans are not a priority.

Structural adjustment policies for privatisation, deregulation and opening of national economies, promoted by international financing and development organisations seem contradictory to the establishment of regulations and controls inherent to mitigation policies.

Many countries view climate change issues as an additional burden which could affect and threaten their competitiveness in the international markets. The creation of commercial blocks together with the globalisation trends on trade put big pressures on the developing countries exporting capability.

In many countries there are problems of population fixation as the geographical asymmetries in development cause migrations to urban areas that are already under strain. Conditions to settlement must then be promoted in the less favoured areas.

If no action is taken, the countries miss opportunities and resources made available after the FCCC.

Internal strengths

The existence alternatives to fossil fuels in energy production or the replacement of imported fuels by less polluting national alternatives.

The existence of national industries that could benefit from mitigation strategies or that require only moderate and realisable technological improvements. This can be seen as an opportunity for development.

Strengths can be identified through the analysis of positive side effects of emission reduction strategies that would offset some of the incremental costs of implementing those strategies. The positive side effects refer to: the *negative cost potential*, namely reductions caused by options that are cheaper than the technologies currently in use; the *economic double dividend*, such as the possible positive effects on growth and employment; and the *environmental double dividend*, namely the correlation between ghg emission reductions and the alleviation of other environmental nuisances such as acidification, urban congestion or deforestation.

Internal weaknesses

The lack of necessary technologies, know-how, or resources for implementing mitigation options, together with insufficient qualified manpower are weaknesses and reasons for the state of already inefficient and ill-managed energy systems in developing countries, the insufficient regional coordination and cooperation on resources utilisation and energy planning issues.

Other weaknesses can be identified by the presence of conflicts between possible mitigation strategies and sectoral policies. Examples of such conflicts are: land use for agriculture versus land use for afforestation; use of indigenous energy resources versus the use of clean and

eventually imported fuels; use of more efficient but imported technologies versus use of less efficient but locally produced technologies; higher prices for energy versus the need for keeping lower, in some cases subsidised, prices as a mechanism of income distribution among the population or for increasing comparative advantages for exports.

When discussing the circumstances of developing countries (Section 4) we described other conditions and limitations that mitigation studies meet. Those have to do with the planning environment, the planning praxis, and the institutional set-up.

After the construction of a matrix for a particular country with opportunities, strengths, threats, and weaknesses including items related to the issues mentioned above, possible strategies could be:

Mini-mini strategies

Attack organisational problems in the current institutional set-up that might originate loss of opportunities for getting external resources. Check the information structure and competencies.

Study the possibility of changing agriculture in areas prone to migration to more labour and land intensive cultures, optimising the use of existing agriculture soils and fixing population.

Mini-maxi strategies

Study the possibility of joint actions that could improve the technological stand of some energy-related industries. Use resources to study the possible non-regret options that could result in economic benefits with fast investment return.

Use international help to introduce energy saving into educational programmes. Seek opportunities for assistance in technical training and improved educational programmes at the highest technical education institutions, for example by inviting foreign experts.

Introduce a long term national mitigation programme.

Maxi-mini strategies

Study mitigation options that bring double dividend in the form of improved competitiveness. Explore local resources as part of a program for population settlement and fixation.

Maxi-maxi strategies

Consider the possibilities for use of indigenous alternative fuels and energy sources. Support industry programmes that result in higher energy efficiency. Study the extent in which national industry services and products can contribute to mitigation actions. Sponsor contacts between industry and foreign partners that allow for joint ventures and technology transfer.

Conjugate national action on acidification and water decontamination with greenhouse gas emissions abatement.

This process should be followed by a selection or prioritisation of strategies and by discussing further implementation aspects [8].

6.2 A Mixed Approach

Suppose a situation where the national Climate Change Commission or government decided to start and provide resources for a particular mitigation study, for example a study on cost of greenhouse gas abatement. This decision might have germinated from a strategic phase like the classical approach described in Section 6.1. To be as fruitful as possible this study should include the concerns of decision makers and other partners. This might result in some limitations on the options included by the analysts, as well as in the inclusion of factors they would not otherwise take into account. On the other side, the decision makers will have a stronger interest on the study and a better commitment to the study progress and results. This mixed approach can be used, for example, as part of an initial consensus workshop.

The decision makers, namely governmental partners, have a global knowledge of the country systems and perceive problems from a top national point of view. They are not interested in the mechanisms of climate change, but only in consequences, possibilities, risks, and demands that climate change issues might cause. For them, the ideal way of decision making includes the control of the environment, the knowledge of the problem, and the knowledge of the different alternatives. They prefer objective criteria according to their policies, and choose the best solution in a rationalistic manner. It is impossible to meet those requisites when dealing with climate change because of the complexity and dimensions of the problem. On the other side, the team elaborating the study is composed of people who are more used to deal with details in their own fields. They know the economics of abatement and the technical possibilities. They might lack, however, a grip of national reality, meaning that the options included in their assessment would often be impracticable or politically unacceptable, because of factors that they do not consider in their analysis, but that decision makers use as eliminatory or important criteria.

Borrowing one analogy from Etzioni [4], one may see the political decision makers using a wide angle lens and the study working team employing a telephoto lens. The wide angle does not allow detail, whereas the telephoto is too narrow for allowing to construct the whole picture within the available time. The picture resulting from this study should be a mixture of these two views, giving the overall picture and presenting more detail where it seems more interesting and useful. This mixed approach to strategic development of responses to climate change tries to accomplish this demand.

Although the mixed approach brings actors together to discuss climate change and talk about specific aspects of it, the discussion cannot elaborate on technical details and the group must be prepared for that. Disagreement on effects of relevant measures might give leads to aspects that need further investigation. In order to take advantage of this approach, the participants should be introduced to any methods used and prepare some short homework, as for example a list of the mitigation options and ideas they think most relevant in their context.

The goal of this approach is not simply to find mitigation options and strategies, but rather to make a kind of initial focusing or screening:

- Discuss the feasibility of all mitigation options proposed;
- Come up with new options;
- Enrich the consciousness to the problems with new aspects; and
- Trigger evolutions in attitude and opinion in the participants.

6.2.1 Example of using the mixed approach

The tool used in this example is again the SWOT matrix. In this case SWOT appears with the role of discussion pathfinder and the final result is not a set of strategies but some principles, importance relations, and directives. The use of a method already familiar to some of the participants may be advantageous.

We suggest a general procedure, illustrated in Figure 4, for idea generating to be used as a part of the building of a SWOT matrix. The principle is that participants in the meeting or workshop come with their ideas on mitigation issues, options, scenarios, or whatever they are concerned with, and from those the planning process takes shape. Suppose that a participant considers a certain mitigation option interesting. Then the group can discuss requisites of that kind of option and from that identify strengths and weaknesses. In a similar manner, opportunities and threats can be identified through the expected consequences.

When doing a SWOT analysis there is no rigidity in which of the four classes to place an issue. Sometimes some participants will suggest that *A* is an opportunity and others that it is a threat, while others think of *A* as a strength; should it be considered as a strength and prioritised later, for example? Should several scenarios for evaluating different evolutions, possibilities, or hypothesis be considered? Likewise in the procedure sketched in Figure 4, where requisites could also generate opportunities and threats for example. A facilitator could assist the group in reaching a formulation they are able to agree upon.

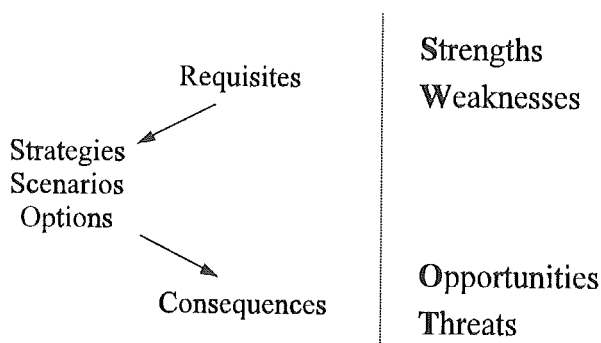


Figure 4 - A procedure for SWOT generation

This short example considers only mitigation in the energy sector. It addresses a fictitious country, though the issues discussed are present in different developing countries (see for example [16]). Participants in a meeting identified requisites and consequences of possible abatement measures:

Development of natural gas supply system

Natural gas is the cleanest of the fossil fuels that are under consideration and so its use would have a strong impact in mitigation. With natural gas, the energy use for cooking could be made more efficient (some would say that the food also tastes better while others point out that allergies to open fires are rising). The risk for accident exists and is aggravated by deficiencies in maintenance. Natural gas demands a network system that must be built and is very expensive, especially in urban areas. It also demands locations for storage. In this hypothetical country, a reserve of natural gas was recently identified and several uses will be considered: transport and distribution of gas to the major towns; production of electricity close to the source; export to two alternative neighbouring countries; and use of gas as feedstock to petrochemical plants.

Increase the efficiency of existing systems by repowering and improvements in transmission and distribution systems

The impact of these kinds of measures in terms of emissions reduction depends on how the energy is generated. Net economical benefits are evident since one will be able to produce more energy at the same total cost. The reasons for the current inefficiencies lie in lack of managing skills in the energy system. There is also the need for a network for servicing and technical parts. Other maintenance problems exist which would be alleviated by better management. Some generating centres have been stopped, on average about one month per year, because of these inefficiencies. The electricity generation system consists in a few hydroelectric power plants along two rivers plus some old thermoelectric plants running on coal and oil. There are also smaller generators scattered around more isolated areas.

Advanced coal power generation technologies such as pulverised coal combustion, atmospheric and pressurised bed combustion, advanced combustion turbines, combined cycle.

The use of newer technologies allows for increased efficiency and less coal usage bringing much lower emissions. These may still cause some local nuisance through noise and other impacts. In addition to the normal effect of rising the temperature of the water used to cool the generator, these technologies also pollute the water used to wash the coal, which in turn should be treated. They can use the existent sources of coal, but technology and know-how are necessary to treat it, wash it, pulverise it, etc. These options demand huge investments and the facilities must be built from scratch.

Business as usual

If the situation is allowed to evolve without interference caused by efficiency concerns and other mitigation issues, the investment in capacity will have to be higher in the future. The government considers building two new power plants over the next six to eight years. These plants are planned to have three power generating units of 200 MW each but it would be better if part of the investments could be delayed, which is difficult under the current conditions and

demand trends. The debt of the public energy company represents already about twenty percent of the national debt. Current trends also show the threat of higher dependence on imported fuel. Environmental impacts of the current technology are accumulating, namely through acidification and water contamination, which might have a nasty impact on fragile fluvial ecosystems.

This option is the one where the decision makers actually do not have to make mitigation decisions, but apparently only adapt to the situations that might come by, which can be very negative, but not until later. It also does not demand extra investment in the short term. The life quality in some regions is not expected to increase under this scenario.

Efficiency improvement on household electrical appliances.

In some electrical appliances large efficiency improvements are attainable. The effect in the electricity bill of a household that buys a cheaper to run (more energy efficient) appliance allows for recovering investments in between two and four years. The economic benefit is not only on the user side, but also on the state side because of an increased ability to delay investments in energy generation. In this country, many homes do not yet use electrical appliances, though the situation is evolving. There exists local manufacture of electrical household appliances but it produces inefficient products, so imports would have to increase. This action has a disperse character and demands an information campaign that alerts the population to the advantages of using more efficient appliances. The success of such a campaign is doubtful since most households cannot afford the additional initial investment.

Optimise the methane release from coal mining.

This option consists in improving the recovery of gas that is released inside coal mines. Traditionally, this operation was done for safety reasons. The government has indications that the activity in the existing coal mines is expected to decrease in the next few years.

As mentioned above, the elements in the SWOT matrix are not so much included in order to build strategies, but for helping the decision process. Many ideas come to the table while building the matrix rather than at the end of a final analysis. Remember that some links between the items already exist due to the process employed in building the matrix. Table 1 presents some internal and external factors used in the discussion and used to build the SWOT matrix. For guidance of the study working team and other partners about the relevance of the factors in the matrix, some of them could be classified as more important, persistent, dominant, or imperative. Table 1 condenses this information by marking those factors with a dotted underline.

Opportunities	Threats
<p>Meet abatement potential</p> <p><u>Increase energy generation efficiency</u></p> <p><u>Use indigenous fuels</u></p> <p><u>Explore natural gas</u></p> <p><u>Attain net benefits without investment in facilities</u></p> <p>Reduce use of coal</p> <p>Obtain benefits on the end user side</p> <p><u>Secure technology transfer</u></p> <p>Improve safety in mining</p> <p>Energy self sufficiency</p>	<p><u>Risk of accidents</u></p> <p>Investments don't pay back</p> <p>Local nuisance caused by new plants</p> <p><u>Water contamination</u></p> <p><u>Increase in foreign debt</u></p> <p>Dependence on imported fuel</p> <p>Other environmental impacts</p> <p><u>Deterioration of life quality</u></p> <p>Increased environmental impacts</p>
Strengths	Weaknesses
<p>Available ground for construction</p> <p><u>Available natural gas reserve</u></p> <p>Available sources of coal</p> <p><u>Experience with different technologies</u></p> <p><u>Local manufacture of appliances</u></p>	<p>Deficiencies in maintenance</p> <p>Lack of capital for investment</p> <p>No network for gas distribution</p> <p><u>Poor management</u></p> <p><u>No network for servicing and parts</u></p> <p>Low quality energy supply</p> <p><u>Locally made appliances inefficient</u></p> <p>Low information dissemination capability</p>

Table 1 - Elements of the SWOT matrix, including some ranking information

After the analysis and discussion, some of the options initially considered by the study team were dropped or reformulated. For example, the idea of changing to more efficient household appliances was dropped, as the necessary internal requirements are mostly weaknesses. However, under the same option, consideration of the particular appliances of air conditioners, lightbulbs, and refrigeration units seemed interesting and the government would be willing to support industry in the modernisation process. That option should be examined closer and be included in the cost studies.

The distribution of natural gas to households was ruled out in the medium term because of the investment needed. Selling gas to a neighbour country is interesting but optimism should be refrained since the main prospect buyer has currently a surplus generating capacity and is undergoing a severe crisis.

In terms of efficiency improvement in energy generation, the impact of liberalisation and deregulation should be assessed, also with respect to the exploration of renewable and new energy sources. Natural gas reserves could be included. The government sees deregulation and

liberalisation as a possibility for "automatic" improvement of efficiency in energy production and distribution. The technology planned for one of the projected power plants rates too high in the greenhouse gas emissions and should be re-assessed.

The participants did not reach agreement on the role or evolution of the domestic electricity market. The agreement was that several scenarios would be considered, with different evolutions of the current asymmetries.

By using this SWOT matrix at a later stage, the team elaborating a study will be able to have an approximate idea on whether to use resources on examining other mitigation options that are brought later to their attention. This can be done by analysing how well they "score" in the framework built in this phase through Table 1.

6.3 Comparison

Both approaches try to structure the way climate change issues are addressed by planners and decision makers.

The classical approach is placed closer to decision making and might be risky if it decides implementation measures at investment level for example, because of the lack of detail. On the other hand it improves the current situation in guiding the coordination of responses to climate change. Many examples exist, where the planning at the decision making level was too strong and caused waste of resources [4].

The mixed approach is a step that brings into the decision process factors that complement the limitation of the top-down approach. The mixed approach brings the views of the decision makers and of the planners together, improving the understanding of the technical aspects by the first improving the understanding of the country reality and circumstances by the second. This approach might demand that some participants, namely the more technical, do and present some preliminary research, for example a preliminary list of mitigation options. Such an approach would favour the country studies on greenhouse gas abatement by attacking some of their bigger shortcomings.

7. Conclusions

Climate change and abatement of greenhouse gas emissions belong to the most complex problems humanity has ever tried to deal with. Current attempts to address them at national level suffer not only from that complexity but also face other obstacles due to the particular circumstances in those countries.

The approaches used in the existing country studies are not especially adequate to deal with the time horizon or the uncertainties involved. They often lack expression of organisational and implementation issues and consideration of criteria that the decision makers would select as pertinent. This entails a contradiction since only if government representatives have a detailed understanding of what is at stake when designing mitigation strategies and greenhouse gas inventories and projections can they hope to hold their own with developed country negotiators

under negotiations following the FCCC [11]. It is then twofold important to bring together the decision makers and the people dealing with the technical study details and overcome the institutional problems by identifying competencies.

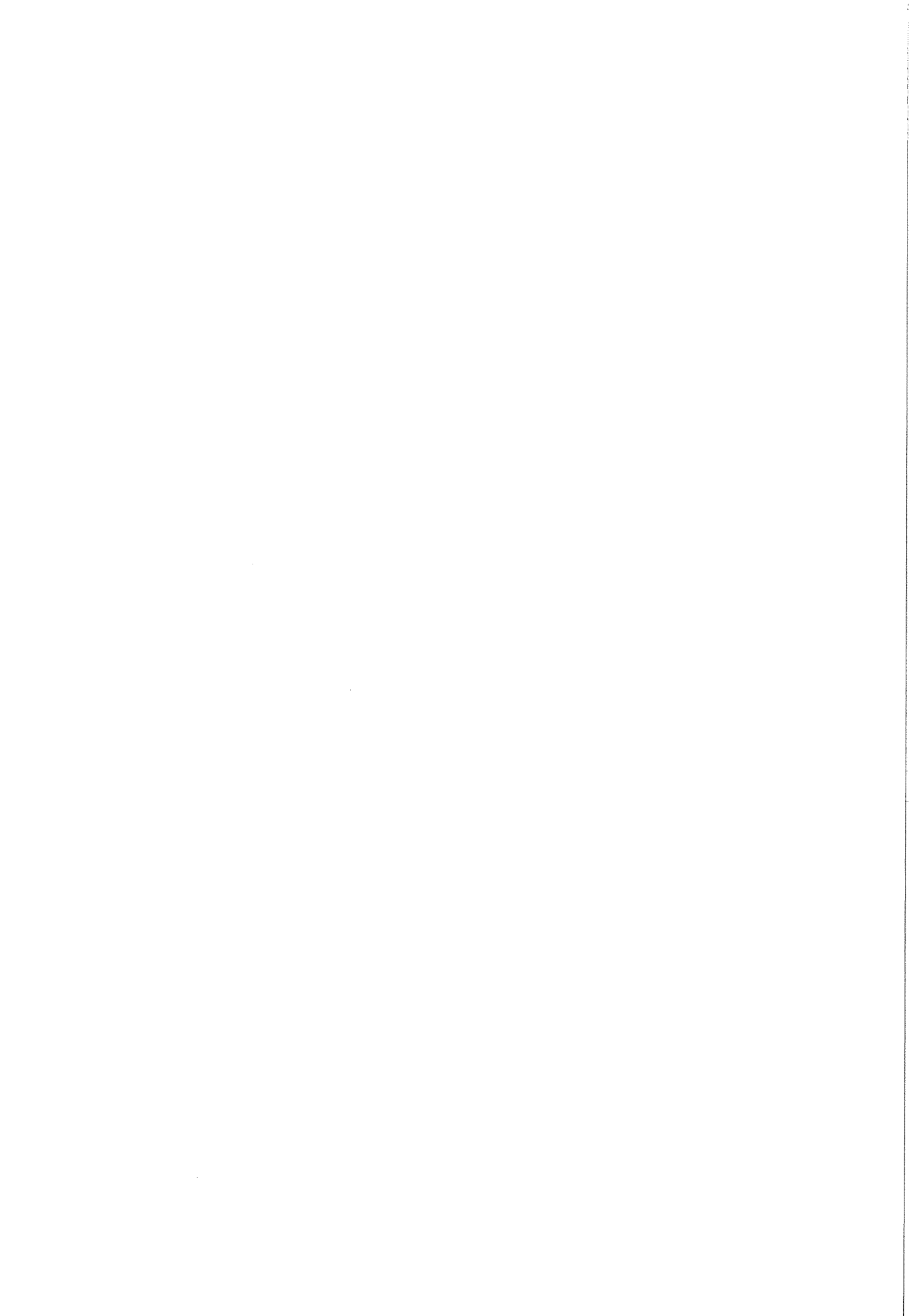
One should differentiate and adapt to the situation of different countries, combining development with mitigation. The designation "developing countries" ranges from countries extremely poor, striving for survival under famine and war environments and where climate change is the least of the issues, to countries that have relatively high development indexes. Dealing with climate change in the way described in this paper, assumes that some organisation capability exists and that thinking about term strategies is possible in the country.

An important task is gathering an institutional set-up (as suggested here with the climate change commission) that is able to act as a climate change policy advisor and ensure continuity in addressing climate change related issues. Priority must also be given to information channels, namely between the climate change commission and the decision making and political spheres.

The issues involved in the strategic planning of responses to climate change suggest using problem structuring methods in the planning activities. This was exemplified by examples of two approaches to address two different planning situations. A classical approach can be used for the planning of high-level strategies, while a mixed approach is suggested for a mid-level planning. The first is important for supporting the design of country strategies, while the second is dedicated to win over the limitations identified in current country studies. Both these approaches make use of the SWOT matrix, but with different purposes.

References

- [1] Dyson, R., *Strategic Planning: Models and Analytical Techniques*, Wiley (1990).
- [2] Emery, J.C., *Organizational Planning and Control Systems*, London (1979).
- [3] *Environment and Energy ministry, Miljoindikatorer 1995*, Copenhagen, Danish environment and energy ministry (1995) (in Danish).
- [4] Etzioni, A., *Mixed-scanning: a 'third' approach to decision-making*, Public Administration Review, December (1967), 385-392.
- [5] Merkus, H., *The Framework Convention on Climate Change: some thoughts on joint implementation*, Ministry of Housing, Physical Planning and Environment, CCD/Paper 11, The Netherlands (1992).
- [6] Mintzberg, H., *The Rise and Fall of Strategic Planning*, Prentice Hall (1994).
- [7] Nilsson, Annika, *Greenhouse Earth*, Wiley (1992).
- [8] Olesen, T., Vival, R.V.V., *TOWS: metode og anvendelser*, Center for Tele-Information, Technical University of Denmark (1995) (in Danish).
- [9] Rosenhead, J. (ed), *Rational analysis for a problematic world*, Wiley (1989).
- [10] Rosenhead, J., *What's the Problem? An introduction to problem structuring methods*, Interfaces 26 (1996) 117-131.
- [11] Speth, J.G., *Climate change & development: A challenge of commitment*, United nations climate change bulletin 8 (1995).
- [12] UNEP, *UNEP Greenhouse Gas Abatement Costing Studies - Phase Two*, Risø (1994).
- [13] United Nations, *United Nations Framework on Climate Change*, New York (1992).
- [14] United Nations, *Kyoto Protocol to the United Nations Framework on Climate Change*, (1997).
- [15] Vidal, R.V.V., *POT: Planning, Organisation and Tele-Information*, Center for Tele-Information, Technical University of Denmark (1996).
- [16] Villavicencio, Arturo, *Mitigation assessment: an approach for screening and evaluating options*, manuscript, UNEP Collaborating Centre on Energy and Environment, Riso, Denmark (1995).
- [17] Weihrich, Heinz, *The TWOS matrix: A tool for situational analysis*, Long Range Planning 15 (1982) 54-66.



ESTIMAÇÃO DO λ ÓTIMO PARA PREVISÃO DA VOLATILIDADE ATRAVÉS DA METODOLOGIA RISKMETRICS™ PARA ATIVOS DO MERCADO FINANCEIRO BRASILEIRO

Lacir J. Soares

Marcelo C. Medeiros

Dept. Engenharia Elétrica
Pontifícia Univ. Católica do Rio de Janeiro
Rua Marquês de São Vicente, 255 - Gávea - 22543-900
Rio de Janeiro - RJ - Brasil

Abstract

The goal of this work is to investigate the returns of the TELEBRÁS PN asset, negotiated at the São Paulo stock market, and to fit two different volatility models: RiskMetrics™ and GARCH(1,1).

Resumo

Este trabalho tem por objetivo investigar a série de retornos da ação TELEBRÁS PN, cotada na bolsa de valores de São Paulo, Brasil, e estimar dois modelos de volatilidade para os dados: RiskMetrics™ e GARCH(1,1).

Keywords

Forecasting, time-series, finance.

1. Introdução

Nos modelos lineares a variância condicional de uma série temporal é suposta invariante no tempo. Entretanto, evidências empíricas, principalmente na literatura econométrica, apontam para existência de dados onde a variância condicional depende da informação disponível. O estudo da volatilidade, isto é, a variação da variância condicional de uma série temporal, é muito importante para a análise de risco no mercado financeiro.

Este trabalho tem o objetivo de estudar a série de retornos TELEBRÁS PN, cotada na bolsa de valores de São Paulo, Brasil. Inicialmente foi feita uma análise dos dados através de estatísticas descritivas e eliminação dos possíveis efeitos auto-regressivos e dos dias da semana na evolução temporal da série. Em seguida, dois modelos de volatilidade foram ajustados e comparados: o GARCH(1,1) e a metodologia RiskMetrics™.

1.1 RiskMetrics™

O RiskMetrics™ é um conjunto de ferramentas que permite a profissionais do mercado financeiro estimar a exposição do mercado ao risco dentro do que é chamado base *Value-at-Risk* (VaR). O RiskMetrics™ possui três componentes básicos:

- um conjunto de metodologias para a medida de risco de mercado;
- conjunto de dados de volatilidade e correlação usado no cálculo do risco de mercado;
- sistemas de *software*, onde os métodos para estimação do risco de mercado estão implementados.

Este trabalho tem por objetivo utilizar as ferramentas do RiskMetrics™ para estimação da volatilidade de ativos financeiros.

Na metodologia RiskMetrics™ a previsão de volatilidade 1-passo à frente é calculada através de uma média móvel exponencialmente amortecida dos retornos passados:

$$\sigma_{t+1|t} = \sqrt{(1-\lambda) \sum_{i=1}^{i=T} \lambda^i r_{t-i}^2}, \quad (1)$$

onde r_i é o retorno do ativo no instante t .

Supondo um número infinito de observações dos retornos, a equação (1) pode ser reescrita em uma forma recursiva, dada pela equação.

$$\sigma_{t+1|t} = \sqrt{\lambda \sigma_{t|t-1}^2 + (1-\lambda)r_t^2} \quad (2)$$

1.2 Os Modelos ARCH e GARCH

Desde a introdução do ARCH (*Autoregressive Conditional Heterokedasticity*) por Engle, (1982), variações, extensões e aplicações desta classe de modelos têm sido desenvolvidas a uma taxa cada vez maior. O estudo de modelos condicionalmente heterocedásticos foi motivado principalmente pelo interesse na modelagem da volatilidade de séries temporais provenientes do mercado financeiro.

O artigo original de Engle verificava se a variância da inflação do Reino Unido era maior em certos períodos do que em outros. Nesta aplicação, havia a necessidade de separar os movimentos previsíveis (média) de movimentos imprevisíveis (resíduo). O modelo ARCH faz com que a variância dos resíduos seja previsível.

Engle, Lilien & Robins, (1987) propuseram uma extensão onde a média condicional da série é influenciada pela variância condicional, o ARCH-M, ou ARCH na média.

A generalização do modelo ARCH, onde a variância condicional é uma média móvel dos resíduos ao quadrado e dos valores passados da variância condicional, foi proposto por Bollerslev, (1986). Este modelo foi chamado de GARCH (*Generalized Autoregressive Conditional Heterokedasticity*) e tem sido vastamente utilizado em aplicações reais.

Outras especificações da família ARCH foram desenvolvidas, dentre as quais, a proposta do GARCH exponencial, ou EGARCH (Nelson, 1991), merece destaque. Esta classe de

modelos difere do GARCH usual em diversas formas; a mais importante é que observações negativas da série implicam em maior volatilidade do que observações positivas. Este facto pode ser confirmado estudando-se retornos de ativos financeiros.

Harvey, Ruiz & Shepard, (1994) propuseram uma classe promissora de modelos chamados de Volatilidade Estocástica, onde a volatilidade é tratada como um processo estocástico não observável e tem analogia direta com os processos de difusão a tempo contínuo utilizados em finanças.

1.2.1 O ARCH(p)

Sejam y_t e I_t , respectivamente, o valor de uma série temporal e a informação disponível no instante t . A média e a variância condicional são definidas, respectivamente, por:

$$m_t = E[y_t | I_t] \quad (3)$$

e

$$h_t = E[(y_t - m_t)^2 | I_t], \quad (4)$$

onde m_t e h_t são funções das componentes de I_t . Se h_t não for constante, a série é heterocedástica.

O modelo ARCH(p) é definido por:

$$\begin{aligned} y_t &= h_t^{1/2} \varepsilon_t, \\ h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i}^2, \end{aligned} \quad (5)$$

onde ε_t é um ruído branco normalmente distribuído de média nula e variância unitária ($\varepsilon_t \sim N(0,1)$).

O modelo ARCH(p) será estacionário, bem definido, se $\alpha_0 > 0$ e $\alpha_1 \dots \alpha_p \geq 0$. Os parâmetros do modelo da equação (5) são estimados pelo método da máxima verosimilhança. Maiores detalhes podem ser encontrados em Engle, (1982).

1.2.2 O GARCH(p,q)

O modelo ARCH(p) proposta por Bollerslev, (1986) é definida por:

$$\begin{aligned} y_t &= h_t^{1/2} \varepsilon_t, \\ h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}, \end{aligned} \quad (6)$$

onde $\varepsilon_t \sim N(0,1)$ e

$$\begin{aligned} p &\geq 0, \quad q > 0, \\ \alpha_0 &> 0, \quad \alpha_i \geq 0, \quad i = 1, \dots, q, \\ \beta_i &\geq 0, \quad i = 1, \dots, p. \end{aligned}$$

O modelo GARCH(p,q) é estacionário, se e somente se, $\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i < 1$. Da mesma

forma que o modelo anterior, os parâmetros desconhecidos são estimados por máxima verosimilhança.

2. Descrição dos Dados

Neste trabalho foram utilizados os retornos diários simples da TELEBRAS PN (de 30/06/94 até 30/12/96), totalizando 617 observações.

Antes de ajustar os modelos de volatilidade diretamente nos retornos, foram analisados os possíveis efeitos existentes na sua média condicional (*ajustes para a média*), em particular, os dias da semana e autocorrelação serial. Este é o procedimento padrão adotado nos trabalhos de estimação da volatilidade de ativos. Inicialmente a série original de retornos foi regredida em variáveis *dummies* representando os dias da semana. Caso uma ou mais *dummies* apresentem efeito significativo, os resíduos desta regressão serão utilizados numa segunda regressão, na qual procura-se ajustar um possível efeito de autocorrelação serial, estimando um modelo AR(p). Na ocorrência de efeitos autoregressivos significantes, os resíduos da segunda regressão serão considerados como sendo os retornos efetivos, que serão utilizados nos modelos de volatilidade. Como as séries são heterocedásticas utilizamos a correção de White (White, 1990) no cálculo do erro padrão de estimação das regressões. Os resultados podem ser encontrados na tabela 1. A tabela 2 apresenta as estatísticas descritivas para a série temporal.

$$r_t = -0.0030 (0.0024) \times \text{SEG} + 0.0013 (0.0031) \times \text{TER} + 0.0021 (0.0027) \times \text{QUA} + 0.0017 (0.0032) \times \text{QUI} + 0.0033 (0.0039) \times \text{SEX} + 0.0020 (0.0048) \times \text{VESP} + 0.0012 (0.0065) \times \text{APOS} + v_t$$

$$r_t = 0.0011 (0.0010) - 0.0200 (0.0762) \times r_{t-1} - 0.0413 (0.0817) \times r_{t-2} - 0.0714 (0.0733) \times r_{t-3} - 0.0384 (0.0625) \times r_{t-4} - 0.1023 (0.0570) \times r_{t-5} - 0.0998 (0.0755) \times r_{t-6} + \varepsilon_t$$

Tabela 1 - Ajustes para média (erro padrão entre parênteses) dos retornos da TELEBRÁS PN

Estatísticas Descritivas	
Número de observações	617
Média	0.004
Desvio padrão	0.044
Variância	0.002
Coefficiente de assimetria	0.410
Coefficiente de curtose	7.997

Tabela 2 - Estatísticas descritivas dos retornos diários da TELEBRÁS PN

Não foi encontrado nenhum efeito significativo de dias da semana nem autoregressivo nos retornos da TELEBRÁS PN.

Os retornos da TELEBRÁS PN no referido período podem ser vistos na figura 1.

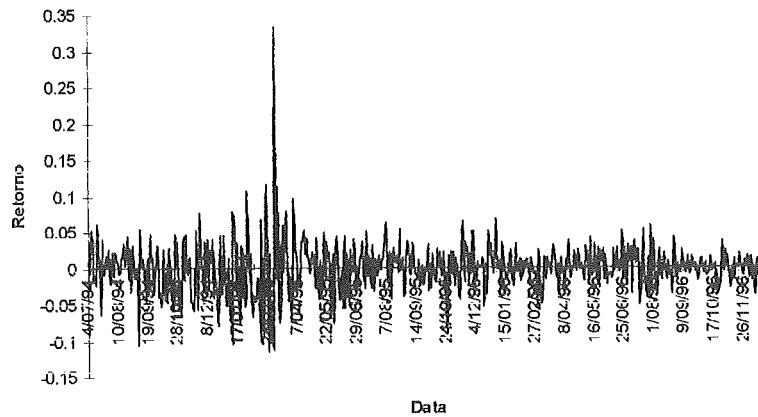


Figura 1 - Retornos TELEBRÁS PN

3. Métodos de Estimação do λ ótimo

O λ ótimo da metodologia RiskMetrics™ foi estimado e os resultados foram comparados com a estimação da volatilidade calculada através de um modelo GARCH(1,1) normal.

Para a estimação dos modelos e para o teste da capacidade de previsão de ambos, a série de retornos foi dividida em duas partes: a primeira contendo 417 observações foi utilizada para a estimação dos parâmetros e a segunda parte, contendo 200 observações, foi utilizada para testar a generabilidade dos modelos.

Para a escolha do fator de descontos ótimo (λ), foram utilizados três métodos: negativo da verossimilhança normal (LKLHD), raiz quadrada do erro quadrático médio (RMSE - *Root Mean Squared Error*) e o erro absoluto médio (MAE - *Mean Absolute Error*). As fórmulas para cada um dos métodos encontram-se a seguir.

$$\text{LKLHD} = \sum_{t=1}^T \left\{ \ln[\sigma_{t|t-1}(\lambda)] + \frac{1}{2} \left(\frac{r_t^2}{\sigma_{t|t-1}^2(\lambda)} \right) \right\} \quad (7)$$

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |r_{t+2}^2 - \sigma_{t+1|t}| \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (|r_{t+1}| - \sigma_{t+1|t})^2} \quad (9)$$

O modelo GARCH(1,1) foi estimado pelo método de máxima verossimilhança.

Para comparar os RiskMetrics com o GARCH(1,1) foram utilizadas as figuras de mérito descritas pelas equações 7, 8 e 9 na fase de previsão.

4. Resultados

4.1 Estimação do λ ótimo

Os resultados para a escolha do λ ótimo estão ilustrados na próxima tabela.

Série	LKLHD	MAE	RMSE
TELEBRÁS PN	0.9090	0.8600	0.8440

Tabela 3 - Valores ótimos de λ

Analisando os valores ótimos encontrados para o λ , percebe-se que é dado um peso maior à volatilidade no instante anterior do que aos retornos passados.

4.2 Estimação do GARCH(1,1)

A tabela 3 apresenta os resultados encontrados para a estimação do GARCH(1,1).

$r_t = h_t \varepsilon_t$		
$h_t = \gamma + \alpha r_{t-1}^2 + \beta h_{t-1}$		
Coefficiente	Valor estimado	Estatística-t
γ	6.49×10^{-5}	2.0665
α	0.16172	2.4450
β	0.7855	13.7941

Tabela 4 - Coeficientes estimados para o GARCH(1,1)

4.3 Previsão da Volatilidade e Comparação com o GARCH(1,1)

A segunda parte da série de retornos da TELEBRÁS PN foi utilizada para testar a capacidade de previsão do modelo proposto pelo RiskMetrics e comparamos os resultados com a previsão obtida por um GARCH(1,1).

Modelo	LKLHD	MAE	RMSE
GARCH(1,1)	-689.8168	4.6718×10^{-4}	0.0148
RiskMetrics ($\lambda = 0.9090$)	-699.6116	3.5967×10^{-4}	0.0123
RiskMetrics ($\lambda = 0.8600$)	-698.0460	3.5379×10^{-4}	0.0122
RiskMetrics ($\lambda = 0.8440$)	-697.1995	3.5297×10^{-4}	0.0122

Tabela 5 - Estatísticas para a previsão da volatilidade

Pela análise dos resultados acima pode-se concluir que a modelagem pelo RiskMetrics foi melhor que pelo GARCH(1,1).

Os próximos gráficos mostram a volatilidade estimada e prevista, pelo GARCH(1,1) e pelo RiskMetrics para os três valores de λ , versus os retornos ao quadrado. Percebemos que todos os dois modelos conseguem capturar eficientemente a dinâmica dos dados.

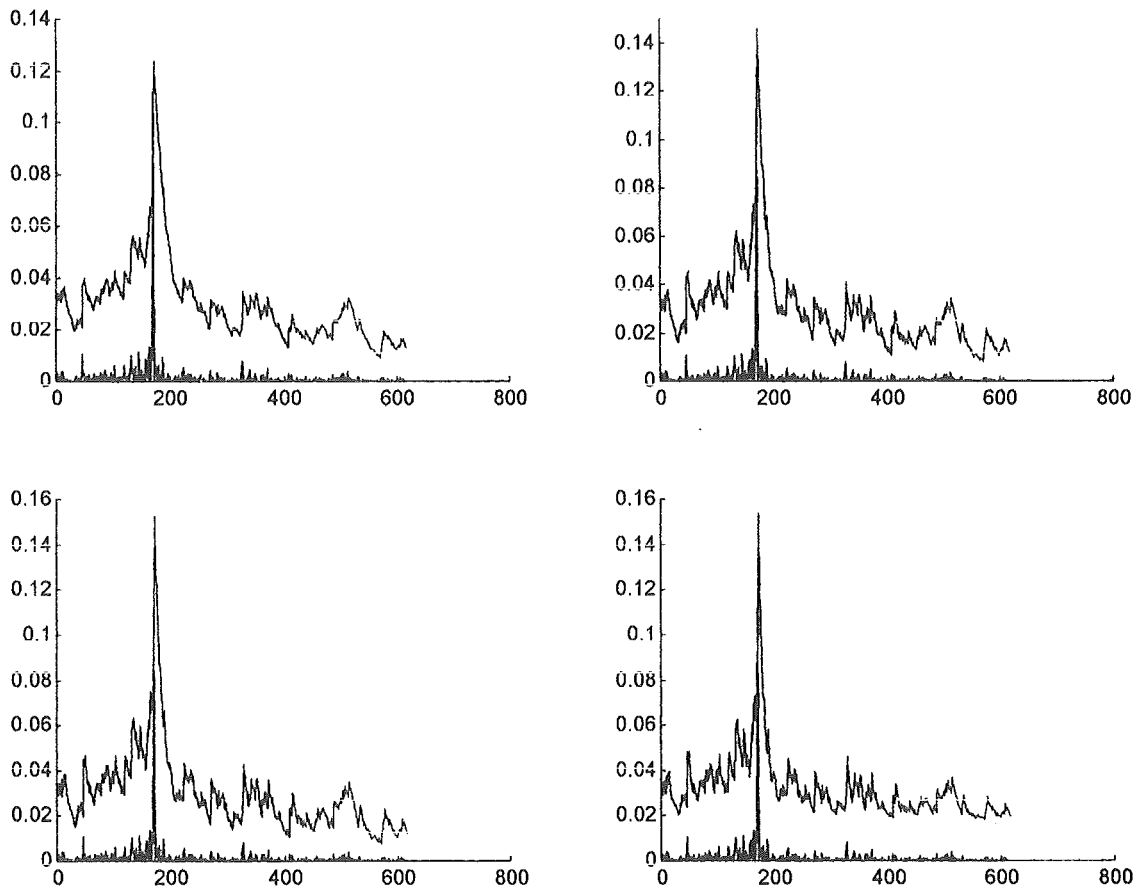


Figura 2 - Volatilidade estimada (-), volatilidade prevista (--) e retornos ao quadrado

Seria interessante uma comparação para mais séries de retornos do mercado financeiro brasileiro e confrontar os resultados com os obtidos para ativos do mercado internacional de capitais. Para uma comparação com outros modelos de volatilidade veja Soares et al., 1997 e Veiga et al., 1997.

5. Referências

- [1] Bollerslev, T.P., *Generalized Autoregressive Conditional Heterocedasticity*, *Journal of Econometrics* 31 (1986) 309-28.
- [2] Engel, R.F., Lilien, D.M. and Robins, R.P., *Estimating Time Varying Risk Premia Term Structure: The ARCH-M Model*, *Econometrica* 55 (1987) 391-408.
- [3] Engle, R.F., *Autoregressive Conditional Heterocedasticity with Estimates of the Variance of UK Inflation*, *Econometric Reviews* 5 (1982) 1-50.
- [4] Engle, R.F., *ARCH Selected Readings*, Oxford University Press (1995).
- [5] Harvey, A.C., Ruiz, E. and Shepard, N.G., *Multivariate Stochastic Variance Models*, *Review of Econometric Studies* 61 (1994) 247-264.
- [6] Nelson, D.B., *Conditional Heterokedasticity in Asset Returns: A New Approach*, *Econometrica* 59 (1991) 347-370.
- [7] *RiskMetrics - Technical Document*, Morgan Guaranty Trust Company - 4th Edition (1996).
- [8] Soares, L.J., Oliveira, F. and Veiga, A., *Comparação de Modelos de Volatilidade Aplicados a Ativos do Mercado Acionário Brasileiro*, *Anais do XXIX Simpósio Brasileiro de Pesquisa Operacional*, Salvador - BA (1997).
- [9] Veiga, A., Medeiros, M. and Fernandes, C., *State Space ARCH: Forecasting Volatility with a Stochastic Coefficient Model*, *Anais do I Computational Finance (V Neural Networks in the Capital Markets)*, Londres-Inglaterra (1997).
- [10] White, H., *A Heterokedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heterokedasticity*, *Econometrica* 48 (1990) 817-38.

INSTRUÇÕES AOS AUTORES

Os autores que desejam submeter um artigo à Investigação Operacional devem enviar três cópias desse trabalho para:

Prof. Joaquim J. Júdice
Departamento de Matemática
Universidade de Coimbra
3000 Coimbra, Portugal

Os artigos devem ser escritos em Português ou Inglês. A primeira página deve conter a seguinte informação:

- Título do artigo
 - Autor(es) e instituição(ões) a que pertence(em)
 - Abstract (em inglês)
 - Resumo
 - Keywords (em inglês)
 - Título abreviado

As figuras devem aparecer em separado de modo a poderem ser reduzidas e fotocopiadas. As referências devem ser numeradas consecutivamente e aparecer por ordem alfabética de acordo com os seguintes formatos:

Artigos: autor(es), título, título e número da revista (livro com indicação dos editores), ano, páginas.

Livros: autor(es), título, editorial, local de edição, ano.

ÍNDICE

P. Carrasco, M. Vaz Pato, Uma heurística neuronal para a elaboração de horários escolares	101
J. P. Costa, J. Clímaco, Um método de programação linear multiobjectivo baseado na projecção dos gradientes das funções objectivo	119
J. L. H. Matias, E. M. G. P. Fernandes, A importância da pesquisa unidimensional na eficiência dos métodos de optimização	137
G. Di Fatta, G. Lo Re, Multicast routing tree optimization	155
S. Kumar, F. Scarmozzino, Characteristics and applications of random search methods for mathematical programming	167
P. Borges, L. Sorensen, A. Villavicencio, R. V. V. Vidal, Strategic approaches to climate change at country level - Focusing on greenhouse gas abatement	183
L. J. Soares, M. C. Medeiros, Estimacão do lamda ótimo para previsão da volatilidade através da metodologia Riskmetrics para ativos do mercado financeiro brasileiro	207

