REFLEXÕES SOBRE TEORIA E PRÁTICA EM PROGRAMAÇÃO NÃO LINEAR COM ESPECIAL ÊNFASE EM UM MÉTODO DO TIPO LAGRANGIANO AUMENTADO

José Mario Martínez

www.ime.unicamp.br/∼martinez

UNICAMP, Brazil

Florianópolis, Setembro 2007

Problema de Programação Não-Linear

PNL

$$\text{Minimizar } f(x)$$

sujeita a

$$h(x) = 0, g(x) \leq 0,$$

$$x \in \Omega,$$

onde $x \in \mathbb{R}^n$, $h(x) \in \mathbb{R}^m, g(x) \in \mathbb{R}^p$.

Importância do PNL

- PNL é um problema matemático
- Com muitas aplicações práticas
- Tomada de Decisões
- Ajuste de Modelos

Algoritmos

- Não há "Métodos Diretos" para resolver PNL.
- PNL se resolve usando Métodos Iterativos
- Um Método Iterativo gera uma sequência de pontos $x^k \in \mathbb{R}^n$ que converge (ou não) a uma solução do problema.
- Métodos Iterativos são programados e implementados em computadores, nos quais, em vez de operações matemáticas "verdadeiras" se fazem operações em ponto flutuante.

“Teoria” e “Prática”

- A Teoria é necessária porque evita fazer infinitos experimentos.
- Teoria útil deve ser capaz de prever o que acontece em muitos experimentos.
- A Teoria normal não se refere às sequências verdadeiras geradas no computador, mas às sequências teóricas definidas pelos algoritmos.
- A analogia entre “sequências verdadeiras” e “sequências teóricas” não é perfeita.
- Procuramos teoria que seja tão preditiva quanto possível.

Estudo de Caso: O Lagrangiano Aumentado PHR

Lagrangiano Aumentado PHR

$$L_\rho(x, \lambda, \mu) = f(x) + \frac{\rho}{2}\left[\left\| h(x) + \frac{\lambda}{\rho} \right\|^2 + \left\| \left(g(x) + \frac{\mu}{\rho} \right)_+ \right\|^2\right]$$

$$(a_+ = \max\{0, a\}, \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p_+)$$

Algoritmo básico baseado em PHR

- Iteração Externa

  "Minimizar" $L_\rho(x, \lambda, \mu)$ sujeita a $x \in \Omega$

- Atualizar multiplicadores λ, μ e parâmetro de penalidade ρ
- Repetir

Dois tipos de Castigo em PHR

Princípio do Castigo por Penalidade (ρ)

O Castigo deve ser Proporcional à Violação das Restrições

$$L_\rho(x,\lambda,\mu) = f(x) + \frac{\rho}{2}\left[\left\|h(x) + \frac{\lambda}{\rho}\right\|^2 + \left\|\left(g(x) + \frac{\mu}{\rho}\right)_+\right\|^2\right]$$

Exagerando nas exigências

Melhor que aumentar a penalidade é fingir que a tolerância à violação a restrições era mais estrita do que na realidade era. Uso de deslocamentos (shifts).

$$L_\rho(x,\lambda,\mu) = f(x) + \frac{\rho}{2}\left[\left\|h(x) + \frac{\lambda}{\rho}\right\|^2 + \left\|\left(g(x) + \frac{\mu}{\rho}\right)_+\right\|^2\right]$$

Virtudes do Lagrangiano Aumentado

Explorar estrutura de subproblemas simples

O conjunto "simples" Ω é arbitrário.
Os métodos de Lagrangiano Aumentado resolvem, sequencialmente, estes problemas simples. O progresso na análise e implementação de algoritmos para resolução de problemas simples produz um efeito positivo quase imediato na qualidade de algoritmos de Lagrangiano Aumentado. Por exemplo, a Otimização em caixas é uma área dinâmica da otimização prática, da qual se podem esperar avanços práticos nos métodos de LA.

Virtudes do Lagrangiano Aumentado

Minimização Global

A minimização global dos subproblemas implica convergência a minimizadores globais do método de Lagrangiano Aumentado.
A Minimização Global em Caixas ($\Omega =$ Caixa) é difícil (como toda minimização global) mas há muita coisa feita que pode ser aproveitada.

Virtudes do Lagrangiano Aumentado

Minimização Global na prática

A maioria dos métodos de minimização em caixas garantem convergência a pontos estacionários. Entretanto bons métodos fazem muito mais do que isso.

Passos de Extrapolação

Passos Mágicos

Aumentam a chance de convergência a minimizadores globais dos subproblemas e, em consequência, do problema original.

Virtudes, virtudes

Quando a Suavidade é duvidosa

A teoria de convergência a minimizadores globais do LA não depende da diferenciabilidade das funções que definem o problema. (Mas sim da continuidade!)

Isto implica que, em situações em que a suavidade é duvidosa (por exemplo pela presença de raízes ou cálculos iterativos internos, etc) o efeito na convergência do LA é benigno.

Mais virtudes

Lagrangiano Aumentado sem Derivadas

O Método de Lagrangiano Aumentado pode ser adaptado ao caso em que as derivadas da função ou restrições não estão disponíveis.
Basta usar um método sem-derivadas para resolver os subproblemas com restrições simples.
As propriedades teóricas de convergência são preservadas.

Mais:

Estrutura de esparsidade variável

Em muitos problemas práticos, a matriz "Hessiana da Lagrangiana", apesar de ser estruturalmente densa (qualquer elemento pode ser diferente de zero) é quase sempre esparsa (dado um elemento do domínio, essa matriz tem poucos elementos diferentes de zero.

Isso é uma dor de cabeça para métodos Newtonianos, que trabalham fatorando, em todas as iterações, a "Matriz do sistema KKT", cuja parte Superior-Esquerda é a Hessiana da Lagrangiana. Entretanto, essa não é uma dificuldade séria quando se usa o Lagrangiano Aumentado apelando para um método "Matrix-free" para resolver os subproblemas.

Ainda mais virtudes

Sistema KKT mal estruturado

Independentemente da densidade da Hessiana do Lagrangiano, a estrutura do sistema KKT pode ser muito desfavorável para fatorações de matrizes esparsas. Isto é um inconveniente sério para métodos newtonianos mas não para implementações do Lagrangiano Aumentado, se o subproblema se resolve usando procedimentos que somente usam produtos de matriz por vetor.

Outra virtude

Problemas com muitas restrições de desigualdade

As restrições de desigualdade são reformuladas, na maioria dos métodos newtonianos, acrescentando variáveis de folga e limitantes inferiores.

Com efeito, a restrição $g_i(x) \leq 0$ é equivalente a $g_i(x) + z_i = 0, \ z_i \geq 0$.

Se o PNL tem muitas restrições de desigualdade, os sistemas que precisam ser resolvidos em cada iteração de um método newtoniano podem ter dimensões astronomicas.

Existem, no contexto newtoniano, procedimentos para aliviar este inconveniente, mas a versão PHR do Lagrangiano Aumentado fornece um remédio radical: Não usar variáveis de folga em absoluto e pagar o preço da descontinuidade das derivadas segundas nos subproblemas.

LA com restrições arbitrárias no nível inferior

Início

$k \leftarrow 1,\ \|V^0\| = \infty,\ \gamma > 1 > \tau,\ \lambda^1 \in \mathbb{R}^m, \mu^1 \in \mathbb{R}^p_+$.

Passo 1: Resolução do Subproblema

Calcular $x^k \in \mathbb{R}^n$, solução aproximada de

$$\text{Minimizar } L_{\rho_k}(x, \lambda^k, \mu^k) \text{ sujeita a } x \in \Omega.$$

Passo 2: Atualizar parâmetro de penalidade e multiplicadores

Definir $V_i^k = \max\left\{g_i(x^k), -\frac{\mu_i^k}{\rho_k}\right\}$.

Se $\max\{\|h(x^k)\|_\infty, \|V^k\|_\infty\} \le \tau\ \max\{\|h(x^{k-1})\|_\infty, \|V^{k-1}\|_\infty\}$,
definir $\rho_{k+1} = \rho_k$. Se não, $\rho_{k+1} = \gamma\rho_k$.
Calcular $\lambda^{k+1} \in [\lambda_{\min}, \lambda_{\max}]^m$, $\mu^{k+1} \in [0, \mu_{\max}]^p$.
Atualizar $k \leftarrow k + 1$ e voltar ao Passo 1.

Convergência a Minimizadores Globais

Teorema

Suponhamos que a região admissível do PNL não é vazia e que cada subproblema é considerado aproximadamente resolvido quando se encontra $x^k \in \Omega$ tal que

$$L_{\rho_k}(x^k, \lambda^k, \mu^k) \leq L_{\rho_k}(y, \lambda^k, \mu^k) + \varepsilon_k$$

para todo $y \in \Omega$, onde $\{\varepsilon_k\}$ é uma sequência de números não-negativos que converge a $\varepsilon \geq 0$.

Então, todo ponto limite x^* da sequência $\{x^k\}$ é admissível e satisfaz:

$$f(x^*) \leq f(y) + \varepsilon$$

para todo ponto admissível y.

Forma geral das Restrições Ω

$$\Omega = \{x \in \mathbb{R}^n \mid \underline{h}(x) = 0, \underline{g}(x) \leq 0\},$$

Minimizadores locais aproximados do Subproblema

Condições ε_k- KKT para o subproblema

$$\|\nabla L_{\rho_k}(x^k, \lambda^k, \mu^k) + \sum_{i=1}^{m} v_i^k \nabla \underline{h}_i(x^k) + \sum_{i=1}^{p} u_i^k \nabla \underline{g}_i(x^k)\| \leq \varepsilon_k,$$

$$u_i^k \geq 0, \underline{g}_i(x^k) \leq \varepsilon_k \text{ for all } i,$$

$$\underline{g}_i(x^k) < -\varepsilon_k \Rightarrow u_i^k = 0 \text{ for all } i,$$

$$\|\underline{h}(x^k)\| \leq \varepsilon_k.$$

Estimativas dos Multiplicadores de Lagrange

Para Restrições de Igualdade

$$\lambda_i^{k+1} = \max\{\lambda_{\min}, \min\{\lambda_i^k + \rho_k h_i(x^k), \lambda_{\max}\}\}$$

Para restrições de Desigualdade

$$\mu_i^{k+1} = \max\{0, \min\{\mu_i^k + \rho_k g_i(x^k), \mu_{\max}\}\}.$$

Dependência Linear Positiva

Dependência Linear Positiva dos Gradientes de Restrições Ativas

Suponhamos que o conjunto admissível de um PNL é dado por $\overline{h}(x) = 0, \overline{g}(x) \leq 0$. Seja $I(x)$ o conjunto de índices das restrições de desigualdade ativas no ponto admissível x. Suponhamos que $I_1 \subset \{1, \dots, \overline{m}\}$, $I_2 \subset I(x)$.

O conjunto de gradiendes das restrições que correspondem aos índices $I_1 \cup I_2$ se diz Positivamente LD se existem coeficientes λ, μ tais que

$$\sum_{i \in I_1} \lambda_i \nabla \overline{h}_i(x) + \sum_{i \in I_2} \mu_i \nabla \overline{g}_i(x) = 0,$$

com $\mu_i \geq 0$ para todo $i \in I_2$ and $\sum_{i \in I_1} |\lambda_i| + \sum_{i \in I_2} \mu_i > 0$.

Em caso contrário, dizemos que estes gradientes são Positivamente LI

Condições de Qualificação

Regularidade (LICQ)

Os gradientes das restrições ativas são LI.

MAIS FORTE (OU SEJA, MAIS RESTRITIVA), QUE:

Mangasarian-Fromovitz

Os gradientes das restrições ativas são positivamente LI.

MAIS FORTE (OU SEJA, MAIS RESTRITIVA), QUE:

CPLD

Dependência Linear Positiva Constante(CPLD)

Se um subconjunto de gradientes de restrções ativas é positivamente LD, então, o mesmo subconjunto de gradientes é LD em uma vizinhança do ponto considerado.

(Qi & Wei, Andreani, J.M.M. & Schuverdt)

Convergência a Pontos Admissíveis

Teorema

Seja x^* um ponto limite de $\{x^k\}$. Então, se a sequência de parâmetros de penalidade é limitada, o ponto x^* é admissível. Se não, pelo menos uma das seguintes possibilidades acontece:

(i) x^* é um ponto KKT do problema

$$\text{Minimize } \frac{1}{2}\left[\sum_{i=1}^{m} h_i(x)^2 + \sum_{i=1}^{p}[g_i(x)_+]^2\right] \text{ subject to } x \in \Omega.$$

(ii) x^* não satisfaz a condição CPLD associada a Ω.

Convergência a pontos KKT

Theorem

Suponhamos que x^* é um ponto limite admissível de $\{x^k\}$ e que satisfaz a condição CPLD. Então, x^* é um ponto KKT do problema PNL.

Limitação dos Parâmetros de Penalidade

Condições sob as quais ρ_k é limitado

- $\lim_{k\to\infty} x^k = x^*$ e x^* é admissível.
- LICQ se satisfaz em x^*. $(\Rightarrow KKT)$.
- A matriz Hessiana do Lagrangiano é definida positiva no subespaço ortogonal aos gradientes das restrições ativas.
- $\lambda_i^* \in (\lambda_{\min}, \lambda_{\max})$, $\mu_j^* \in [0, \mu_{\max})$ para todo i, j.
- Para todo i tal que $g_i(x^*) = 0$, temos que $\mu_i^* > 0$. (Complementaridade estrita no nível superior.)
- Existe uma sequência $\eta_k \to 0$ tal que $\varepsilon_k \leq \eta_k \max\{\|h(x^k)\|, \|V^k\|\}$ para todo $k = 0, 1, 2 \ldots$.

Condição de Otimalidade de Segunda Ordem

Condição de Segunda Ordem Fraca (SOC)

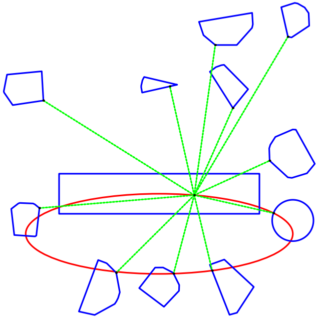
A Hessiana do Lagrangiano é semidefinida positiva no subespaço ortogonal aos gradientes das restrições ativas (chamado "subespaço tangente" de agora em diante).

Teorema

Se os subproblemas são resolvidos com um "método de segunda ordem": Todo ponto limite admissível de $\{x^k\}$, sob uma "condição de qualificação de segunda orden" adequada, satisfaz SOC.

Exemplo de LA com restrições no nível inferior muito estruturadas

Encontrar o ponto no Retângulo mas não na Elipse tal que a soma das distâncias aos Polígonos é mínima.
Restrições do nível superior: (Todos os pontos) $\notin$ Ellipse
Restrições do nível inferior: Ponto Central $\in$ Rectangle, Demais pontos $\in$ Polígonos.



Exemplo

1,567,804 polígonos
3,135,608 variáveis, 1,567,804 restrições do nível superior,
12,833,106 restrições do nível inferior.
Convergência em
10 iterações externas, 56 iterações internas, 133 avaliações de função, 185 segundos.

Razões

Usamos, neste caso, o Método de Gradiente Espectral Projetado (SPG), para resolver os subproblemas. Este método resulta extremamente eficiente devido a que calcular projeções, neste problema, é muito fácil.

ALGENCAN

ALGENCAN é o Algoritmo de Lagrangiano Aumentado PHR onde as restrições do nível inferior $x \in \Omega$ formam uma caixa.

Método para resolver os subproblemas: GENCAN

GENCAN usa:

- Estratégia de Restrições Ativas
- Newton-Inexato dentro das faces
- SPG para abandonar as faces
- Extrapolações e passos mágicos.

"Aspiração Modesta" para ALGENCAN

ALGENCAN deve ser eficiente quando o PNL tem:

- Muitas restrições de Desigualdade, ou
- Estrutura Ruim da Matriz do Sistema KKT.

Exemplo: Problema das Esferas rígidas

Encontrar n_p pontos na esfera unitária de $\mathbb{R}^{n_d}$ maximizando a mínima distância entre pares de pontos.

Formulação PNL

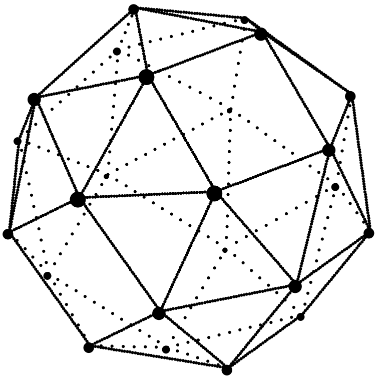
$$
\begin{array}{ll}
\text{Minimize}_{p^i, z} & z \\
\text{subject to} & \|p^i\|^2 = 1, \; i = 1, \ldots, n_p, \\
& \langle p^i, p^j \rangle \leq z, \; i = 1, \ldots, n_p - 1, \; j = i+1, \ldots, n_p,
\end{array}
$$

onde $p^i \in \mathbb{R}^{n_d}$ for all $i = 1, \ldots, n_p$. Este problema tem $n_d \times n_p + 1$ variáveis, n_p restrições de igualdade e

$$n_p \times (n_p - 1)/2$$

restrições de desigualdade.

Esferas Rígidas com $n_d = 3$, $n_p = 24$



Desempenho de ALGENCAN no problema das Esferas Rígidas

Hard-Spheres (3,162)

	Final infeasibility	Final f	Iterations	Time
ALGENCAN	3.7424E-11	9.5889E-01	10	40.15
IPOPT	5.7954E-10	9.5912E-01	944	1701.63

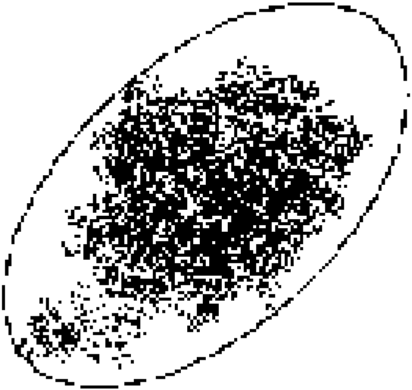
Problema do menor Elipsoide

Encontrar o Elipsoide de menor volume que contem n_p pontos dados em $\mathbb{R}^{n_d}$.

$$\begin{array}{ll} \text{Minimize}_{l_{ij}} & -\sum_{i=1}^{n_d} \log(l_{ii}) \\ \text{subject to} & (p^i)^T L L^T p^i \leq 1, i = 1, \ldots, n_p, \\ & l_{ii} \geq 10^{-16}, i = 1, \ldots, n_d, \end{array}$$

onde $L \in \mathbb{R}^{n_d \times n_d}$ é triangular inferior. O número de variáveis é $n_d \times (n_d + 1)/2$ é o número de restrições de desigualdade é n_p.

Menor Elipsoide cobrindo uma Proteína



Test com Menor Elipsoide

6 variáveis, 20000 restrições de desigualdade.

Enclosing-Ellipsoid (3,20000)

	Final infeasibility	Final f	Iterations	Time
ALGENCAN	8.3449E-09	3.0495E+01	28	1.90
IPOPT	1.1102E-15	3.0495E+01	41	9.45

Bratu

Problema tipo Bratu 3D: Minimizar uma quadrÃtica simples, sujeita à discretização da Equação de Bratu tri-dimensional.

$$\begin{array}{ll} \text{Minimize}_{\, u(i,j,k)} & \sum_{(i,j,k)\in S}[u(i,j,k) - u_*(i,j,k)]^2 \\ \text{subject to} & \phi_\theta(u,i,j,k) = \phi_\theta(u_*,i,j,k), \;\; i,j,k = 2,\ldots,n_p-1, \end{array}$$

onde

$$\phi_\theta(v,i,j,k) = -\Delta v(i,j,k) + \theta e^{v(i,j,k)},$$

e

$$\Delta v(i,j,k) = \frac{v(i\pm 1,j,k) + v(i,j\pm 1,k) + v(i,j,k\pm 1) - 6v(i,j,k)}{h^2},$$

O número de variáveis é n_p^3 e o número de restrições de igualdade é $(n_p-2)^3$. Usamos $\theta = -100$, $h = 1/(n_p-1)$ e $|S| = 7$. Não há restrições de desigualdade.

Resultado de Bratu

$n_p = 20$, $n = 8000$, Restrições: 5832.

Bratu-based $(20, \theta = -100, \#S = 7)$

	Final infeasibility	Final f	Iterations	Time
ALGENCAN	6.5411E-09	2.2907E-17	3	5.12
IPOPT	2.7311E-08	8.2058E-14	5	217.22

Características de Esferas Rígidas, Menor Elipsoide e Bratu

Esferas Rígidas e Menor Elipsoide têm muitas restrições de desigualdade.

Bratu tem uma estrutura complicada na matriz do sistema KKT.

Resumo até aqui

- Apresentamos um método do tipo Lagrangiano Aumentado PHR para Programação Não-Linear.
- Exibimos um conjunto de razões pelas quais a filosofia de Lagrangiano Aumentado nunca deverá ser abandonada, apesar de eventuais modas passageiras.
- Caracterizamos tipos de problemas para os quais o LA-PHR deve ser eficiente.
- Corroboramos com exemplos computacionais a hipótese fundamental sobre a eficiência do método.
- Não afirmamos eficiência nem superioridade para todos os problema possíveis de PNL !!!
- Apresentamos Teoria de Convergência que pretende explicar o que acontece na prática.
- A Teoria está Livre de Pecados no sentido a ser explicitado no resto desta palestra.

Hamartiologia da Otimização: Sete Pecados

Muitas teorias de convergência exibem diferentes tipos de pecados. Peca-se por muitos motivos: Ignorância, Ingenuidade, Necessidade iminente, Indução por parte de uma autoridade superior, Frustração diante do inevitável, etcétera.

- Pecado da Sequência Limitada.
- Pecado da suposição sobre $\varphi(x^k)$.
- Pecado das Hipótesis que não se verificam nunca.
- Pecado das Sequências densas.
- Pecado dos Subproblemas difíceis.
- Pecado das Condições de Otimalidade Duvidosas.
- Pecado de Ocultar os Pecados.

Pecado da Sequência Limitada

Não é pecado supor que a sequência $\{x^k\}$ é limitada, ou a sequência λ^k ou μ^k.
Não é pecado, sempre que você possa dar condições suficientes razoáveis sobre o problema que garantam que a sequência é limitada. Em PNL isto pode não ser fácil (Em Minimização sem restrições é fácil.)
Se você dribla o pecado colocando uma caixa grande artificial, tudo bem, não é pecado, mas nesse caso seu algoritmo deve lidar com caixas, ou seja, as caixas devem formar parte da estrutura do problema.
O perigo é que o teorema Todo ponto de acumulação é solução seja verdadeiro porque, simplesmente, não há pontos de acumulação.

Pecado da Suposição sobre $\varphi(x^k)$

Alguns teoremas pecaminosos dizem "Se a sequência gerada pelo algoritmo satisfaz $\varphi(x^k, \lambda^k, \mu^k, \rho_k) > \varepsilon$ para todo k, então, todo ponto limite é solução".

Isto é muito interessante, mas, a menos que você tenha uma condição suficiente razoável para que a desigualdade $\varphi(x^k, \lambda^k, \mu^k, \rho_k) > \varepsilon$ seja válida, o resultado de convergência é incompleto.

O perigo, de novo, é que seu teorema seja válido simplesmente porque a condição colocada sobre a sequência não se verifica nunca.

Pecado da Sequência Densa

Alguns teoremas irreprensíveis do ponto de vista matemático se baseiam no fato de que, em última instância, o algoritmo pode gerar uma sequência densa no domínio.
Ora, se uma sequência é densa, visita as vizinhanças de todos os pontos, e portanto passa tão perto quanto quiser da solução.
Assim, o Teorema é verdadeiro, e pode afirmar coisas tão mirabolantes como Todo ponto limite é solução global, mas não explica nada.
Ver paper "anónimo" no primeiro número de Mathematical Programming.

Pecado da Tarefa Difícil

Alguns algoritmos incluem a seguinte cláusula: Se [as coisas não vão bem] encontre um ponto admissível.

Os argumentos são irrepreensíveis, mas o algoritmo supõe que "encontrar um ponto admissível" é mais fácil que resolver o PNL original. (Encontrar um ponto admissível pode ser muito difícil, sobretudo quando não há pontos admissíveis.)

Assim, a eficiência teórica destes algoritmos está ligada à eficiência de algum algoritmo associado para encontrar pontos admissíveis.

Dado que encontrar um ponto KKT é, de fato, encontrar um ponto admissível (de outro conjunto de igualdades e desigualdades) a pergunta óbvia é: por que não usar o "outro" algoritmo para resolver o problema original?

Alguns autores, motivados por esta dificuldade, consideram que não vale a pena fazer teoremas de convergência e que tudo é uma questão de bom senso.

Pecado da Condição de Otimalidade Mirabolante

História Imaginária:
Um autor descobre uma condição de otimalidade extraordinária.
Esta condição não usa os gradientes das restrições, o que é surpreendente, pois as condições de Lagrange, como todos sabem, relaciona gradiente de funçào objetivo e gradientes de restrições.
A Nova Condição se verifica quando o minimizador local satisfaz uma nova Condição de Qualificação.
Porém, esta condição de qualificação é, felizmente, muito fraca, ou seja, não-restritiva, ou seja, boa. Tão boa que se verifica em todos os pontos admissíveis com probabilidade 1 !.
O autor introduz um algoritmo para o qual prova que Se um ponto limite satisfaz a condição de qualificação, então satisfaz a condição de otimalidade (como é usual).
O Algoritmo é surpreendentemente simples. Tão simples que nem usa informação sobre as restrições nem sobre seus gradientes.

Condição de Otimalidade Mirabolante

A condição de otimalidade era:

Condição de Otimalidade Mirabolante

A Condição de Otimalidade era:

$$\nabla f(x) = 0$$

Condição de Otimalidade Mirabolante

A Condição de Otimalidade era:

$$\nabla f(x) = 0$$

A Condição de Qualificação era:
O ponto x é interior

Condição de Otimalidade Mirabolante

A Condição de Otimalidade era:

$$\nabla f(x) = 0$$

A Condição de Qualificação era:
O ponto x é interior
O método é:
Alguma variação boba do método do gradiente.

Pecado de Ocultar os Pecados

De fato, este é o único pecado importante.
Resultados fracos, pouco explicativos ou incompletos são admissíveis, sempre que, claro, do ponto de vista matemático sejam corretos.
O único pecado de verdade é pretender que o teorema diz o que não diz, explica o que não explica, esclarece o que não esclarece.

Algumas referências

- R. Andreani, E. G. Birgin, J. M. Martínez and M. L. Schuverdt. On Augmented Lagrangian methods with general lower-level constraints. To appear in *SIAM Journal on Optimization* .
- E. G. Birgin and J. M. Martínez. Improving ultimate convergence of an Augmented Lagrangian method. To appear in *Optimization Methods and Software*.
- R. Andreani, J. M. Martínez and M. L. Schuverdt. On second-order optimality conditions for nonlinear programming. To appear in *Optimization*.
- L. Ferreira-Mendonça, V. L. Lopes and J. M. Martínez. Quasi-Newton acceleration for equality constrained minimization. To appear in *Computational Optimization and Applications*.

- E. G. Birgin and J. M. Martínez. Structured Minimal-Memory Inexact Quasi-Newton method and secant preconditioners. To appear in *Computational Optimization and Applications*.
- R. Andreani, E. G. Birgin, J. M. Martínez and M. L. Schuverdt. Augmented Lagrangian methods under the Constant Positive Linear Dependence constraint qualification. *Mathematical Programming* 111, pp. 5-32 (2008).
- R. Andreani, J. M. Martínez and M. L. Schuverdt, On the relation between the Constant Positive Linear Dependence condition and quasinormality constraint qualification, *Journal of Optimization Theory and Applications* 125, pp. 473–485, 2005.

- E. G. Birgin, C. Floudas and J. M. Martínez. Global Optimization using an Augmented Lagrangian method with variable lower-level constraints.
- M. A. Diniz-Ehrhardt, M. A. Gomes-Ruggiero, J. M. Martínez and S. A. Santos. Augmented Lagrangian algorithms based on the spectral projected gradient for solving nonlinear programming problems. *Journal of Optimization Theory and Applications* 123, pp. 497-517 (2004).