Método do Lagrangiano aumentado

Marina Andretta

ICMC-USP

23 de novembro de 2010

Problema com restrições gerais

Vamos agora resolver problemas de programação não-linear geral

$$\boxed{\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeita a} & h(x) = 0 \\ & g(x) \leq 0 \\ & x \in \Omega, \end{array}} \tag{1}$$

onde

- $\Omega = \{x \in \boldsymbol{R}^n \mid \bar{h}(x) = 0 \text{ e } \bar{g}(x) \leq 0\}$,
- $x \in \boldsymbol{R}^n$, $f : \boldsymbol{R}^n \to \boldsymbol{R}$,
- $h : \boldsymbol{R}^n \to \boldsymbol{R}^m$, $g : \boldsymbol{R}^n \to \boldsymbol{R}^p$,
- $\bar{h} : \boldsymbol{R}^n \to \boldsymbol{R}^{\bar{m}}$, $\bar{g} : \boldsymbol{R}^n \to \boldsymbol{R}^{\bar{p}}$,
- f, h, g, $\bar{h}$, $\bar{g}$ funções suaves.

Método de Lagrangiano aumentado

A função de Lagrangiano aumentado Powell-Hestenes-Rockafellar (PHR) é dada por

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

com $a_+ = \max\{0, a\}$, $\rho > 0$, $\lambda \in \boldsymbol{R}^m$, $\mu \in \boldsymbol{R}^p_+$.

Método de Lagrangiano aumentado

O escalar ρ é chamado de parâmetro de penalidade.

Os vetores λ e μ são os multiplicadores de Lagrange associados às restrições de igualdade e desigualdade, respectivamente.

Note que a função de Lagrangiano aumentado penaliza as restrições h e g que não são satisfeitas pelo ponto x.

Quanto maior o valor do parâmetro ρ, mais penalizadas são as restrições não satisfeitas.

Método de Lagrangiano aumentado: Dados $x_0 \in \boldsymbol{R}^n$, $\lambda^0 \in \boldsymbol{R}^m$, $0 < \mu^0 \in \boldsymbol{R}^p$, escalar $\rho_0 > 0$.

Passo 1: Faça $k \leftarrow 0$.

Passo 2: Se x_k é solução do problema original (1), pare.

Passo 3: Com os valores de λ^k, μ^k e ρ_k fixos, resolva aproximadamente o subproblema

$$\boxed{\begin{array}{ll} \text{Minimizar} & \mathcal{L}_{\rho_k}(x, \lambda^k, \mu^k) \\ \text{sujeita a} & x \in \Omega, \end{array}}$$

Passo 4: Atualize multiplicadores λ^{k+1}, $\mu^{k+1} > 0$ e parâmetro de penalidade $\rho_{k+1} > 0$.

Passo 5: Faça $k \leftarrow k + 1$ e volte para o passo 2.

Método de Lagrangiano aumentado

Note que, no passo 3 do método de Lagrangianos aumentados, é necessário resolver um problema de minimização restrito ao conjunto Ω.

Por isso, Ω é escolhido como um conjunto de restrições *fáceis*.

Quando dizemos que Ω é um conjunto de restrições *fáceis*, queremos dizer que existem métodos eficientes para resolver problemas restritos ao conjunto Ω.

Vamos considerar o caso em que $\Omega = \{x \in \boldsymbol{R}^n \mid \ell \leq x \leq u\}$.

Método de Lagrangiano aumentado

No passo 3, quando dizemos que se deve resolver o subproblema aproximadamente, queremos dizer que, dada uma sequência de $\epsilon_k \to 0$, deve ser usado um método para resolver o subproblema que termine com um ponto x_k tal que

$$\boxed{x_k \in \Omega \text{ e } \|P_\Omega(x_k - \nabla \mathcal{L}_{\rho_k}(x_k, \lambda^k, \mu^k)) - x_k\| \leq \epsilon_k,}$$

com P_Ω projeção ortogonal em Ω.

Atualização dos multiplicadores de Lagrange

Em implementações práticas do método de Lagrangiano aumentado, calculamos

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

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

Estas são estimativas de primeira ordem dos multiplicadores de Lagrange com salvaguardas.

Atualização dos multiplicadores de Lagrange

As salvaguardas definidas por λ_{min}, λ_{max} e μ_{max} ($\lambda_{\text{min}} < \lambda_{\text{max}}$ e $\mu_{\text{max}} > 0$) são necessárias para provar resultados de convergência global.

Sem estas salvaguardas são necessárias hipóteses fortes sobre o problema para garantir a limitação dos parâmetros de penalidade.

Defina

$$\boxed{V_i^k = \max\left\{g_i(x_k), -\frac{\mu_i^k}{\rho_k}\right\}, \ i = 1, \ldots, p.}$$

A atualização do parâmetro de penalidade ρ é feita da seguinte maneira:

- Se $k = 0$, faça $\rho_{k+1} \leftarrow \rho_k$.
- Se $\max\{\|h(x_k)\|_\infty, \|V_i^k\|_\infty\} \leq \tau \max\{\|h(x_{k-1})\|_\infty, \|V_i^{k-1}\|_\infty\}$, com $0 < \tau < 1$, faça $\rho_{k+1} \leftarrow \rho_k$.
- Caso contrário, faça $\rho_{k+1} \leftarrow \gamma\rho_k$, para $\gamma > 1$.

Convergência

Teorema 1: *Seja* $\{x_k\}$ *a sequência gerada pelo método de Lagrangiano aumentado. Seja* x^* *o ponto limite de* $\{x_k\}$*. Então, se a sequência de parâmetros de penalidade* $\{\rho_k\}$ *for limitada, o ponto limite* x^* *é viável. Caso contrário,* x^* *é um ponto KKT do problema*

$$\begin{array}{ll} \textit{Minimizar} & \frac{1}{2}\left[\sum_{i=1}^{m}[h(x)]_i^2 + \sum_{i=1}^{p} \max\{0, [g(x)]_i^2\}\right] \\ \textit{sujeita a} & x \in \Omega. \end{array}$$

Teorema 2: *Seja* $\{x_k\}$ *a sequência gerada pelo método de Lagrangiano aumentado. Suponha que* x^* *seja um ponto limite viável do problema (1). Então,* x^* *é um ponto KKT do problema (1).*

Método de Lagrangiano aumentado

Há várias vantagens no uso de métodos de Lagrangiano aumentado. Algumas delas são:

- Progresso na análise e implementação de métodos de otimização eficientes para resolver problemas simples produzem um efeito positivo quase imediato na eficiência e robustez dos métodos de Lagrangiano aumentado correspondentes.

- Minimização global dos subproblemas implica em convergência a minimizadores globais do método de Lagrangiano aumentado.

Método de Lagrangiano aumentado

- Em vários problemas práticos, a Hessiana do Lagrangiano é uma matriz estruturalmente densa (ou seja, qualquer entrada pode ser diferente de zero para Hessianas calculadas em diferentes pontos) mas geralmente esparsa (dado um ponto no domínio, a Hessiana do Lagrangiano neste ponto é esparsa).

  Métodos Newtonianos geralmente têm dificuldades em problemas deste tipo, tanto em termos de memória como em termos de tempo computacional, pois o padrão de esparsidade muda a cada iteração.

  Esta dificuldade é praticamente irrelevante para métodos do tipo Lagrangiano aumentado que utilizam subalgoritmos que necessitam de pouca memória.

Método de Lagrangiano aumentado

- Independentemente da densidade da Hessiana do Lagrangiano, a estrutura do sistema KKT pode ser muito ruim para fatorações esparsas.

  Este é um problema grave para métodos Newtonianos, mas não para boas implementações de métodos de Lagrangiano aumentado baseados na formulação PHR.

Método de Lagrangiano aumentado

- Se o problema de programação não-linear possui muitas restrições de desigualdade, o uso de variáveis de folga feito por métodos de pontos interiores pode não ser conveniente.

  Há várias maneiras de reduzir os efeitos causados por muitas variáveis de folga, mas eles podem não ser tão efetivos como não usar variáveis de folga.

  A desvantagem de não usar variáveis de folga é a ausência de segundas derivadas contínuas da função de Lagrangiano aumentado. Em muitos casos, isto não se mostra um inconveniente na prática.

Método de Lagrangiano aumentado

- Problemas de grande porte têm a desvantagem óbvia em termos de necessidade de armazenamento.

  Métodos do tipo Lagrangiano aumentado apresentam uma solução: dados do problema podem ser computados apenas quando necessários, usados quando preciso para os subproblemas, sem nenhum armazenamento.

  Isto não é possível quando se usam métodos matriciais, independentemente da estratégia de esparsidade adotada.