Método de Newton modificado

Marina Andretta

ICMC-USP

14 de setembro de 2010

Método de Newton

Como já vimos, o método de Newton com tamanhos de passo unitários converge rapidamente quando se aproxima do minimizador x^*.

No entanto, este algoritmo é inadequado para uso geral, já que ele pode não convergir para uma solução quando o ponto inicial está longe desta. Nestes casos, mesmo que ele convirja, seu comportamento pode ser ruim em regiões nas quais a função não é convexa.

O objetivo ao desenvolver um método prático baseado em Newton é fazê-lo robusto e eficiente em todos os casos.

Método de Newton

Lembre-se que o passo de Newton p_k^N é calculado resolvendo-se o sistema linear simétrico

$$\boxed{\nabla^2 f_k p_k^N = -\nabla f_k.} \tag{1}$$

Para garantir convergência global, pedimos que a direção p_k^N seja de descida, o que acontece quando $\nabla^2 f_k$ é definida positiva.

Método de Newton

Se a Hessiana $\nabla^2 f_k$ não é definida positiva ou está perto de ser singular, a direção p_k^N pode ser de subida ou pode ser excessivamente longa.

Para contornar este problema usaremos duas estratégias:

1. Modificar a matriz Hessiana $\nabla^2 f_k$ antes ou durante a resolução do sistema (1) para que ela se torne suficientemente definida positiva. Este método é chamado de método de Newton modificado.

2. Resolver o sistema (1) usando um método de gradientes conjugados que pára quando uma direção de curvatura negativa é encontrada. Este método é chamado de método de Newton truncado.

Método de Newton modificado

Em muitas aplicações deseja-se usar um método direto para resolver o sistema newtoniano (1).

Se a Hessiana não é definida positiva ou está perto de ser singular, podemos modificar a matriz antes ou durante a resolução do sistema.

Assim, o passo p_k será solução de um sistema como (1), trocando a Hessiana por uma aproximação definida positiva.

A aproximação da Hessiana é obtida somando-se à Hessiana um múltiplo da matriz identidade ou uma matriz completa.

Método de Newton modificado com busca linear: Dado um ponto inicial x_0 e um escalar $\epsilon > 0$.

Passo 1: Faça $k \leftarrow 0$.

Passo 2: Se $\|\nabla f(x_k)\| \leq \epsilon$, pare com x_k como solução.

Passo 3: Faça $B_k \leftarrow \nabla^2 f(x_k) + E_k$. Se $\nabla^2 f(x_k)$ é suficientemente definida positiva, $E_k \leftarrow 0$. Senão, defina E_k de forma que B_k seja suficientemente definida positiva.

Passo 4: Fatore a matriz B_k e resolva $B_k p_k = -\nabla f(x_k)$.

Passo 5: Faça $x_{k+1} \leftarrow x_k + \alpha_k p_k$, α_k satisfaz condições de Wolfe ou é calculado usando *backtracking* com Armijo.

Passo 6: Faça $k \leftarrow k + 1$ e volte para o Passo 2.

Método de Newton modificado

É possível obter resultados de convergência bons para o Algoritmo 1, dado que a estratégia de escolha de E_k satisfaça a propriedade de fatoração modificada limitada.

Esta propriedade diz que a sequência de matrizes $\{B_k\}$ possuem número de condição limitado sempre que a sequência de matrizes $\{\nabla^2 f(x_k)\}$ é limitada. Ou seja

$$\boxed{cond(B_k) = \|B_k\|\|B_k^{-1}\| \leq C, \quad C > 0, \quad \forall k.} \tag{2}$$

Método de Newton modificado

Teorema 3: *Seja f com derivadas segundas contínuas em um conjunto aberto $\mathcal{D}$. Suponha que o ponto inicial do Algoritmo 1 seja tal que o conjunto de nível $L = \{x \in \mathcal{D} | f(x) \leq f(x_0)\}$ é compacto.*
Então, se a propriedade de fatoração modificada limitada vale, temos que

$$\lim_{k \to \infty} \nabla f(x_k) = 0.$$

Método de Newton modificado

Consideremos agora a ordem de convergência do método de Newton modificado com busca linear. Suponha que a sequência de iterandos x_k convege a um ponto x^* no qual a Hessiana é suficientemente definida positiva (ou seja, $E_k = 0$ para todo k suficientemente grande). Neste caso, temos que $\alpha_k = 1$ para todo k suficientemente grande e o método de reduz ao método de Newton puro. Isso garante convergência quadrática.

Para problemas em que $\nabla^2 f(x^*)$ está perto de ser singular, não há garantia de que E_k se anulará. Neste caso, convergência pode ser apenas linear.

Modificações na Hessiana

Veremos agora algoritmos para garantir que a matriz B_k seja suficientemente definida positiva. Além disso, B_k deve ser bem-condicionada e desejamos somar o mínimo possível à Hessiana para que a informação de segunda ordem seja mantida ao máximo.

Começaremos com uma estratégia ideal, baseada na decomposição em autovalores de $\nabla^2 f(x_k)$

Modificação do autovalor

Considere um problema para o qual, na iteração k, temos $\nabla f(x_k) = (1, -3, 2)$ e $\nabla^2 f(x_k) = diag(10, 3, -1)$, que é indefinida.

Pelo teorema da decomposição espectral, podemos definir $Q = I$ e $\Lambda = diag(\lambda_1, \lambda_2, \lambda_3)$ e escrever

$$\boxed{\nabla^2 f(x_k) = Q\Lambda Q^T = \sum_{i=1}^{n} \lambda_i q_i q_i^T.} \tag{3}$$

Modificação do autovalor

O passo de Newton (solução de (1)) é $p_k^N = (-0.1, 1, 2)$, que não é uma direção de descida, já que $\nabla f(x_k)^T p_k^N = 0.9 > 0$.

Uma ideia para substituir a matriz $\nabla^2 f(x_k)$ do sistema newtoniano por uma matriz B_k definida positiva seria trocar todos os autovalores negativos da matriz Hessiana por um número positivo pequeno δ que seja maior do que a precisão u da máquina, digamos $\delta = \sqrt{u}$.

Modificação do autovalor

Supondo que a precisão da máquina seja 10^{-16}, a matriz resultante neste exemplo seria

$$\boxed{B_k = \sum_{i=1}^{2} \lambda_i q_i q_i^T + \delta q_3 q_3^T = diag(10, 3, 10^{-8}).} \qquad (4)$$

Esta matriz B_k é numericamente definida positiva e tem a curvatura nas direções dos autovetores q_1 e q_2 preservada.

Modificação do autovalor

No entanto, a direção de busca baseada na Hessiana modificada seria

$$\boxed{p_K = -\sum_{i=1}^{2} \frac{1}{\lambda_i} q_i (q_i^T \nabla f_k) - \frac{1}{\delta} q_3 (q_3^T \nabla f_k) \approx -(2 \times 10^8) q_3.} \quad (5)$$

Para valores pequenos de δ, esta direção é quase paralela a q_3 (com pouca contribuição de q_1 e q_2) e muito longa.

Apesar de f decrescer ao longo da direção p_k, seu tamanho viola o espírito do método de Newton, que é baseado na aproximação quadrática da função objetivo que é válida numa vizinhança do ponto atual x_k.

Existem muitas estratégias que podem ser consideradas:

- Trocar os sinais dos autovalores negativos em (3), que seria o mesmo que considerar $\delta = 1$ neste exemplo.
- Definir o último termo de (5) como nulo, para que a direção de curvatura negativa não seja considerada.
- Adaptar a escolha de δ para garantir que o tamanho do passo não seja muito grande.
- Etc.

Apesar do grande número de estratégias possíveis, não há uma que seja considerada ideal.

Modificação do autovalor

A modificação feita em (4) para a matriz do exemplo pode ser considerada ótima no seguinte sentido: se A é uma matriz simétrica com decomposição espectral $A = Q \Lambda Q^T$, então a matriz de correção ΔA de norma de Frobenius mínima que garante que $\lambda_{\min}(A + \Delta A) \geq \delta$ é dada por

$$\boxed{\Delta A = Q diag(\tau_i) Q^T, \quad \text{com } \tau_i = \begin{cases} 0, & \lambda_i \geq \delta, \\ \delta - \lambda_i, & \lambda_i < \delta. \end{cases}} \tag{6}$$

Modificação do autovalor

Note que a matriz ΔA não é diagonal no caso geral.

A matriz modificada, neste caso, passa a ser

$$\boxed{A + \Delta A = A + Q diag(\tau_i) Q^T.}$$

Modificação do autovalor

Usando outra norma para o cálculo de ΔA, podemos obter uma matriz diagonal.

Suponha novamente que A seja uma matriz simétrica com decomposição espectral $A = Q\Lambda Q^T$. A matriz de correção ΔA com norma euclidiana mínima que satisfaz $\lambda_{\min}(A + \Delta A) \geq \delta$ é dada por

$$\boxed{\Delta A = \tau I, \qquad \text{com } \tau = \max\{0, \delta - \lambda_{\min}(A)\}.} \tag{7}$$

Neste caso, a matriz modificada tem a forma

$$\boxed{A + \tau I.}$$

Portanto, todos os autovalores de A são deslocados e todos são maiores do que δ.

Estes resultados mostram que tanto modificações diagonais como não diagonais podem ser consideradas. Apesar de não se saber que estratégia produz melhores resultados, vários métodos para modificar a Hessiana já foram propostos e implementados, produzindo bons resultados.

Soma de um múltiplo da identidade

Talvez a ideia mais simples seja encontrar um escalar $\tau > 0$ tal que $\nabla^2 f(x_k) + \tau I$ seja suficientemente definida positiva.

Sabemos que τ deve satisfazer (7), mas geralmente não temos uma boa estimativa do menor autovalor da Hessiana.

Uma ideia é usar o fato que o maior autovalor da matriz A, em valor absoluto, é limitado pela norma de Frobenius $\|A\|_F$. Isso sugere a seguinte estratégia:

Soma de um múltiplo da identidade

Cholesky com múltiplo da identidade somado: Dada uma matriz A.

Passo 1: Faça $\beta \leftarrow \|A\|_F$, $k \leftarrow 0$.

Passo 2: Se $\min a_{ii} > 0$, então

faça $\tau_0 \leftarrow 0$.

Senão

faça $\tau_0 \leftarrow \beta/2$.

Passo 3: Tente aplicar a fatoração de Cholesky incompleta para obter $GG^T = A + \tau_k I$.

Se a fatoração foi feita com sucesso, pare com G.

Senão, faça $\tau_{k+1} \leftarrow \max(2\tau_k, \beta/2)$, $k \leftarrow k+1$ e volte para o Passo 3.

Soma de um múltiplo da identidade

Esta estratégia é simples e pode ser preferível à fatoração de Cholesky modificada. No entanto, ela possui dois inconvenientes.

1. O valor de τ gerado pelo algoritmo pode ser desnecessariamente grande, o que prejudicaria a direção de Newton modificada, tornando-a muito próxima à direção de máxima descida.

2. Para todo valor de τ_k, é necessário recalcular a fatoração $A + \tau_k I$. Isto pode ser custoso quando vários valores de τ_k são gerados.

Fatoração de Cholesky modificada

Uma estratégia popular usada para modificar a matriz Hessiana que não é definida positiva é realizar a fatoração de Cholesky e aumentar os elementos da diagonal durante a fatoração (quando necessário) para garantir que eles sejam suficientemente positivos.

Esta estratégia de fatoração de Cholesky modificada foi desenvolvida para satisfazer dois objetivos:

1. Garantir que os fatores de Cholesky modificados existam e que são limitados pela norma da Hessiana verdadeira.
2. Não mudar a Hessiana quando esta é suficientemente definida positiva.

Fatoração de Cholesky modificada

Primeiramente, vejamos como é feita a fatoração de Cholesky.

Toda matriz simétrica definida positiva pode ser escrita como

$$\boxed{A = LDL^T,}$$

com L matriz triangular inferior com diagonal unitária e D matriz diagonal com elementos positivos na diagonal.

Fatoração de Cholesky na forma LDL^T: Dada uma matriz A.

Passo 1: Para $j = 1, 2, ..., n$,

faça $c_{jj} \leftarrow a_{jj} - \sum_{s=1}^{j-1} d_s l_{js}^2$,

Faça $d_j \leftarrow c_{jj}$.

Para $i = j + 1, ..., n$,

faça $c_{ij} \leftarrow a_{ij} - \sum_{s=1}^{j-1} d_s l_{is} l_{js}$,

faça $l_{ij} \leftarrow c_{ij} / d_j$.

Pode-se mostrar que os elementos da diagonal d_j são todos positivos quando A é definida positiva.

O algoritmo apresentado difere um pouco dos algoritmos de fatoração de Cholesky tradicionais, que calculam a fatoração $A = GG^T$, com G triangular inferior. De fato, $G = LD^{1/2}$.

Se A é indefinida, a fatoração $A = LDL^T$ pode não existir. Mesmo que exista, o algoritmo apresentado é numericamente instável quando aplicado a estas matrizes, pois os elementos de D e L podem se tornar arbitrariamente grandes.

Assim, a estratégia de calcular a fatoração LDL^T e depois modificar a diagonal para que seus elementos sejam positivos pode não funcionar ou a matriz resultante LDL^T pode ser muito diferente de A.

Fatoração de Cholesky modificada

Em vez disso, a matriz A será modificada durante a fatoração, de modo que todos os elementos de D sejam suficientemente positivos e, então, os elementos de L e D não sejam muito grandes.

Para controlar a qualidade da modificação, são usados dois parâmetros positivos δ e β. Durante o cálculo da j-ésima coluna de L e D, pede-se que que

$$\boxed{d_j \geq \delta, \quad |g_{ij}| \leq \beta, \quad i = j+1, ..., n.} \tag{8}$$

Fatoração de Cholesky modificada

Para satisfazer estes limitantes, precisamos modificar apenas um passo do algoritmo de fatoração de Cholesky: a fórmula para calcular a diagonal d_j é substituída por

$$\boxed{d_j = \max\left\{|c_{jj}|, \left(\frac{\theta_j}{\beta}\right)^2, \delta\right\}, \quad \text{com } \theta_j = \max_{j<i\leq n} |c_{ij}|.}$$

Note que os limitantes (8) valem, já que $c_{ij} = l_{ij} d_j$ e, portanto,

$$|g_{ij}| = |l_{ij}\sqrt{d_j}| = \frac{|c_{ij}|}{\sqrt{d_j}} \leq \frac{|c_{ij}|\beta}{\theta_j} \leq \beta, \quad \forall i > j.$$

Fatoração de Cholesky modificada

Note que θ_j pode ser calculado antes de d_j, pois os elementos c_{ij} do segundo laço do algoritmo de fatoração de Cholesky não envolvem d_j.

Este fato sugere que os cálculos sejam agrupados de modo que c_{ij} seja computado antes dos elementos diagonais d_j.

Para tentar diminuir o tamanho das modificações, permutações simétricas de linhas e colunas são feitas, para que no j-ésimo passo da fatoração, a linha e coluna j são aquelas que possuem o maior elemento da diagonal.

Fatoração de Cholesky modificada: Dados $A \in \boldsymbol{R}^{n\times n}$, $\delta > 0$, $\beta > 0$.

Passo 1: Para $j = 1, 2, ..., n$, faça $c_{kk} \leftarrow a_{kk}$.

Passo 2: Para $j = 1, 2, ..., n$,

encontre o índice q tal que $|c_{qq}| \geq |c_{ii}|$, $i = j, ..., n$.

Permute as linhas e colunas i e q.

Para $s = 1, ..., j-1$, faça $l_{js} \leftarrow c_{js}/d_s$.

Para $i = j+1, ..., n$, faça $c_{ij} \leftarrow a_{ij} - \sum_{s=1}^{j-1} l_{js} c_{is}$.

Faça $\theta_j \leftarrow 0$.

Se $j \leq n$, então $\theta_j \leftarrow \max_{j<i\leq n} |c_{ij}|$.

Faça $d_j \leftarrow \max\{|c_{jj}|, (\theta_j/\beta)^2, \delta\}$.

Se $j < n$ então

para $i = j+1, ..., n$, faça $c_{ii} \leftarrow c_{ii} - c_{ij}^2/d_j$.

Fatoração de Cholesky modificada

O algoritmo para fatoração de Cholesky modificada requer, aproximadamente, $n^3/6$ operações aritméticas, que é quase o mesmo que a fatoração de Cholesky tradicional.

No entanto, as permutações de linhas e colunas fazem com que dados devam ser movimentados no computador, o que pode representar altos custos em problemas de grande porte.

Não é necessário armazenamento além do necessário para A, já que o fator triangular L, a matriz diagonal D e os valores intermediários c_{ij} podem sobrescrever os elementos de A.

Fatoração de Cholesky modificada

Vamos denotar por P a matriz de permutações associada com as permutações de linhas e colunas feitas no Passo 2 do algoritmo de fatoração de Cholesky modificada. Não é difícil notar que o algoritmo produz a fatoração de Cholesky de uma matriz $PAP^T + E$, ou seja,

$$\boxed{PAP^T + E = LDL^T = GG^T,}$$

onde E é uma matriz diagonal não-negativa que é nula se A é suficientemente definida positiva.

Examinando as fórmulas de c_{jj} e d_j, fica claro que os elementos da diagonal de E são $e_j = d_j - c_{jj}$.

Também fica claro que somar e_j a c_{jj} na fatoração é equivalente a somar e_j a a_{jj} da matriz original.

Fica faltando apenas especificar a escolha dos parâmetros δ e β.

A constante δ é geralmente escolhida perto da precisão da máquina u. Uma escolha típica é

$$\boxed{\delta = u \max\{\gamma(A) + \xi(A), 1\},}$$

com $\gamma(A)$ elemento da diagonal de maior magnitude e $\xi(A)$ elemento fora da diagonal de maior magnitude. Ou seja,

$$\boxed{\gamma(A) = \max_{1 \leq i \leq n} |a_{ii}|, \quad \xi(A) = \max_{i \neq j} |a_{ij}|.}$$

Fatoração de Cholesky modificada

Uma sugestão para a escolha de β é

$$\boxed{\beta = \max\left\{\gamma(A), \frac{\xi(A)}{\sqrt{n^2-1}}, u\right\}^{1/2},}$$

que tenta minimizar a norma das modificações $\|E\|_\infty$.

É possível mostrar que a matriz B_k obtida pela aplicação da fatoração de Cholesky modificada à Hessiana $\nabla^2 f(x_k)$ tem número de condição limitado, ou seja, o limitante (2) vale para algum valor de C.