MÉTODOS DE PONTOS INTERIORES PARA PROGRAMAÇÃO QUADRÁTICA

Abel Soares Siqueira
abel.s.siqueira@gmail.com
Bolsista

Francisco A. M. Gomes
chico@ime.unicamp.br
Orientador

DMA - IMECC - UNICAMP

PIBIC/CNPq

Palavras-Chave: Programação Quadrática - Métodos de Pontos Interiores - Otimização

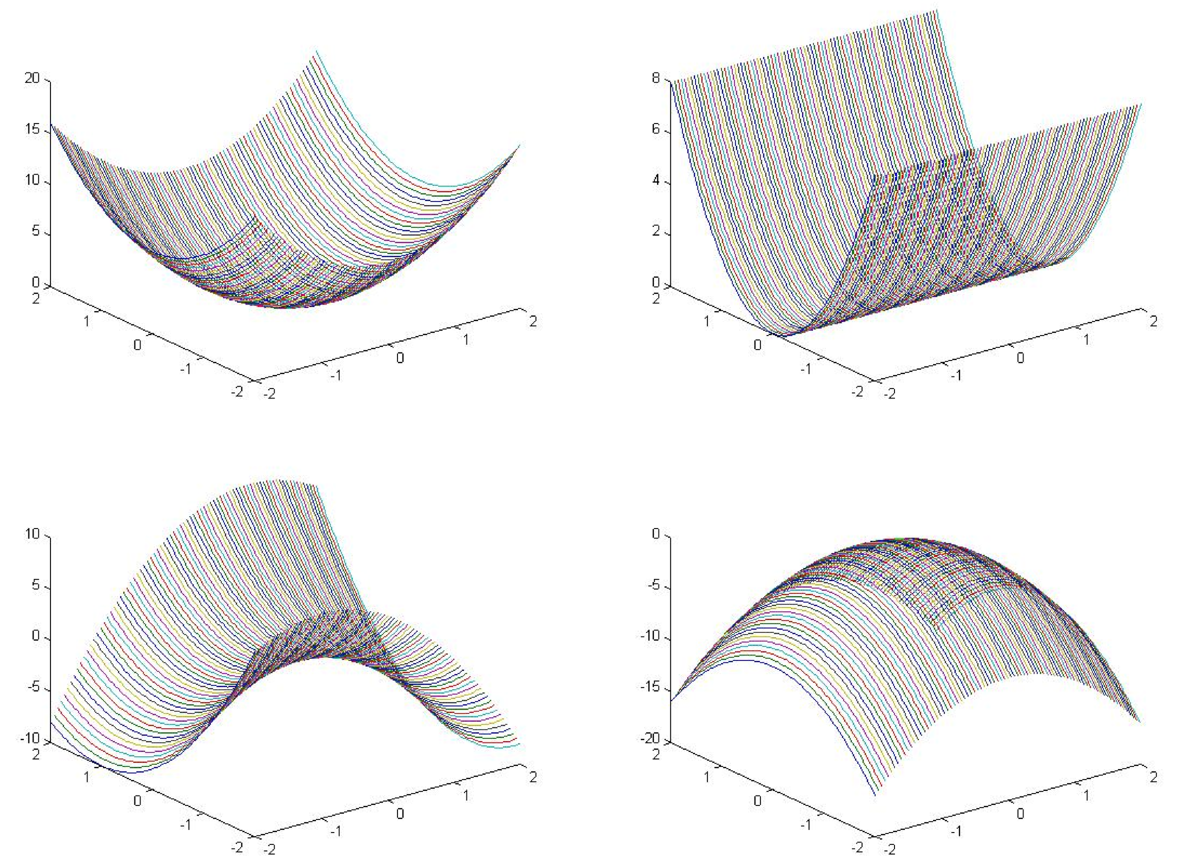
Programação Quadrática

Problemas de programação quadrática são um tipo específico de problema de programação não-linear, onde a função objetivo é uma função quadrática e as restrições são lineares. Um problema de programação quadrática é um problema que pode ser escrito como

$$\min_x q(x) = \frac{1}{2}x^TGx + d^Tx + h, \qquad \text{suj. a } Ax = b, x \geq 0,$$

onde $G \in \mathbb{R}^{n\times n}$ é simétrica, $x, d \in \mathbb{R}^n$ e $h \in \mathbb{R}$, $A \in \mathbb{R}^{m\times n}$, $m < n$ e $b \in \mathbb{R}^m$.

Alguns exemplos de quadráticas no $\mathbb{R}^2$ estão a seguir.



Pontos Interiores

Os métodos de pontos interiores podem ser vistos como métodos que geram uma sequência de pontos que satisfazem as restrições de desigualdade estritamente, onde essa sequência converge para a solução do problema. Uma das maneiras de gerar esses pontos é considerar uma barreira associada as restrições de desigualdade. No problema (1), teríamos o seguinte:

$$\min_x \phi(x,\mu) = \frac{1}{2}x^TGx + x^Td + h - \mu e^T \log x, \text{ suj. a } Ax = b$$

Pela própria definição da função objetivo, a variável x só pode assumir valores estritamente positivos. Dessa maneira, garantimos que todas as iterações geradas pelo método permaneçam positivas.

Um exemplo pode ser visto a seguir com o problema

$$\min_{x,y} f(x,y) = x^2 + (y-1)^2, \qquad \text{suj. a } x, y \geq 0,$$

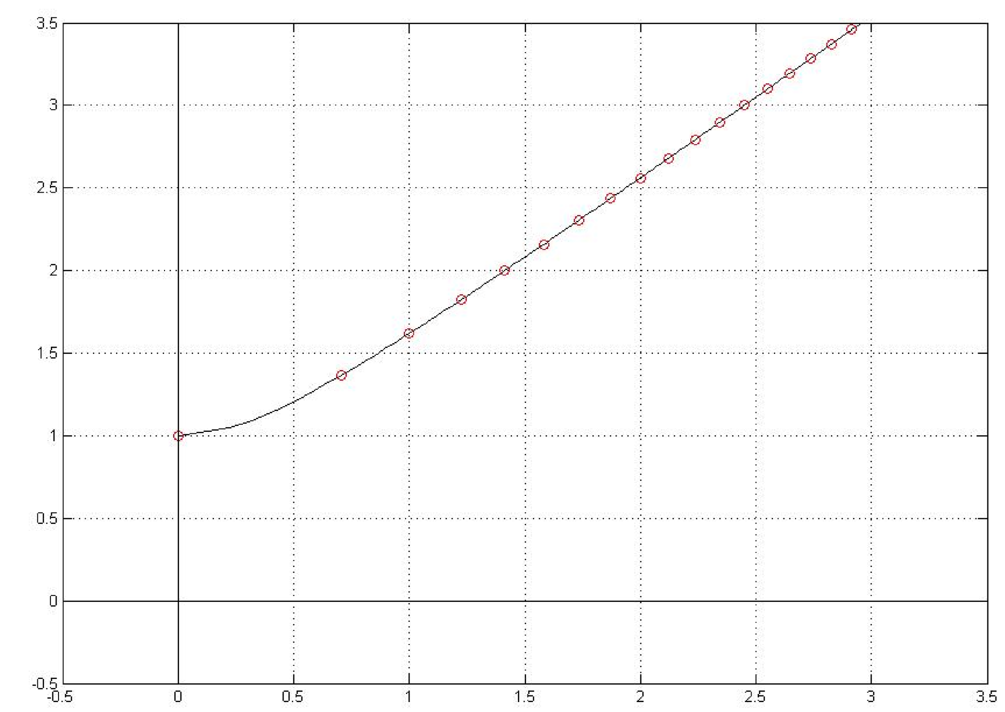
cuja solução podemos ver que é $x = 0$ e $y = 1$. O problema de barreira associado é

$$\min_{x,y} \phi(x,y,\mu) = x^2 + (y-1)^2 - \mu[\log(x) + \log(y)].$$

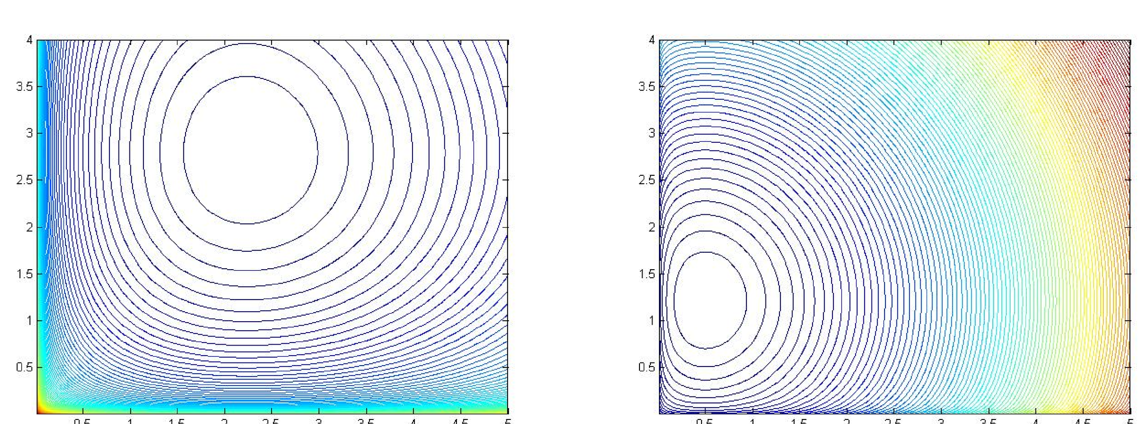
A solução desse problema, em função de μ é

$$x(\mu) = \frac{\sqrt{2\mu}}{2}, \qquad y(\mu) = \frac{1+\sqrt{1+2\mu}}{2}$$

Um gráfico dessas soluções em função de μ está a seguir:



Veja as curvas de contorno de $\phi(x,y,\mu)$ para $\mu = 10$ e $\mu = 0.5$ na figura a seguir:



Espaço Nulo

O espaço nulo de uma matriz $A \in \mathbb{R}^{m\times n}$ é o subespaço definido por

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\}.$$

Dizemos que uma matriz N forma uma base para o espaço nulo de A se as colunas de N são vetores linearmente independentes de $\mathcal{N}(A)$ e se o número de colunas de N é igual a dimensão desse subespaço. Se A tem posto completo, então $N \in \mathbb{R}^{n\times n-m}$. É fácil ver que $AN = 0$.

Podemos agora reduzir o seguinte problema:

$$\min_s q(s) = \frac{1}{2}s^TBs + s^Tg, \qquad \text{suj. a } As = 0.$$

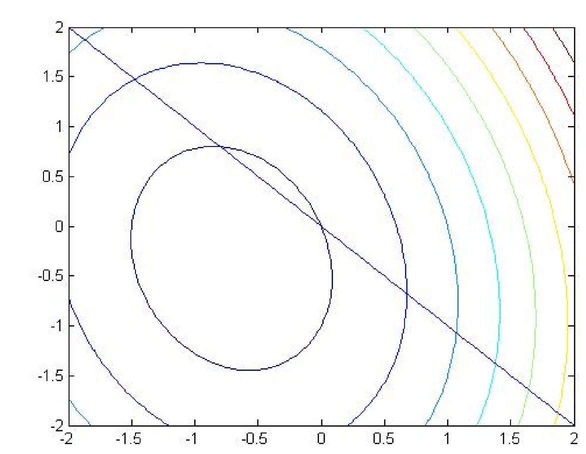
Para isso, fazemos a substituição $s = Nu$, daí o problema se reduz a

$$\min_u \tilde{q}(u) = \frac{1}{2}u^TN^TBNu + u^TN^Tg.$$

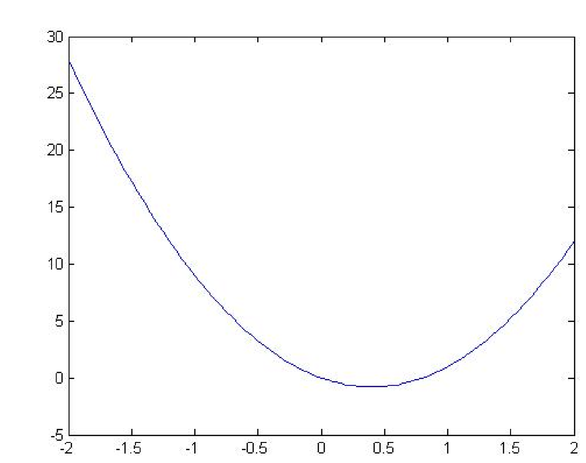
Dessa forma, podemos resolver um problema sem restrições. Como exemplo, considere o problema

$$\min_{x,y} q(x,y) = 2x^2 + xy + 4y^2 + 2x + 6y, \qquad \text{suj. a } x + y = 0$$

O gráfico das curvas de nível do problema é o seguinte:



A reta na figura representa a restrição $Ax = 0$. Fazemos a substituição $(x,y)^T = (1,-1)^Tu$ Dessa maneira, nosso problema se reduz a $\min_u \tilde{q}(u) = 5u^2 - 4u$, cujo gráfico é:



Resolvendo esse problema, poderemos obter a solução original.

Algoritmo

Nosso algoritmo trabalha com dois loops, um interno e um externo. O loop externo é responsável por diminuir o valor de μ até que ele seja suficientemente próximo de 0. O loop interno é responsável por minimizar a função $\phi(x,\mu)$ para um dado μ.

Nosso loop interno parte de um ponto x^0 tal que $Ax^0 = b, x^0 \geq 0$. Em cada iteração, ele calcula um passo s^k tal que $x^{k+1} = x^k + s^k$ e $Ax^{k+1} = b, x^{k+1} > 0$. Dessa relação, notamos que $As^k = 0$. Para calcular esse passo, utilizamos a expansão de taylor de segunda ordem da função $\phi(x,\mu)$ em torno do ponto x^k:

$$\phi(x_k + s, \mu) \approx \phi(x_k,\mu) + s^T\nabla\phi(x_k,\mu) + \frac{1}{2}s^T\nabla^2\phi(x_k,\mu)s.$$

Definimos então $g_k = \nabla\phi(x_k,\mu) = Gx^k + d - \mu(X^k)^{-1}e$ e $B_k = \nabla^2\phi(x_k,\mu) = G + (X^k)^{-1}Z^k$, e minimizamos a função quadrática obtida em uma região definida em torno de x^k. Desse modo, s^k é a solução de

$$\begin{aligned} \min_s \; & q_k(s) = \frac{1}{2}s^TB_ks + s^Tg_k, \\ \text{suj. a } & As = 0 \\ & \|s\| \leq \Delta. \end{aligned}$$

Para resolver esse problema utilizamos uma versão modificada do método dos gradientes conjugados, chamado de método dos gradientes conjugados projetados precondicionado. A principal diferença desse método é a possibilidade de encontrar um ponto estacionário mesmo quando a matriz não for definida positiva.