Exercício Programa 3
Matrizes Ortogonais e o Problema de Quadrados Mínimos

Lucas Magno
7994983

Introdução

Este EP consiste em se resolver um sistema linear sobredeterminado na forma

$$Ax = b$$

onde

$$\begin{aligned} A &\in \mathbb{R}^{n\times m} \\ x &\in \mathbb{R}^{m} \\ b &\in \mathbb{R}^{n} \end{aligned}$$

a fim de minimizar a norma do resíduo, ou, equivalentemente, sua norma ao quadrado, dada por

$$\begin{aligned} \|r\|_2^2 &= \|b - Ax\|_2^2 \\ &= \sum_{i=1}^{n}(b_i - (Ax)_i)^2 \\ &= \sum_{i=1}^{n} r_i^2 \end{aligned}$$

Ou seja, o problema se resume em encontrar x que minimize os r_i^2, o que dá o nome ao Método dos Mínimos Quadrados.

Matrizes Ortogonais

Para tal, vale a pena o estudo de matrizes ortogonais, que são definidas como qualquer matriz Q tal que

$$Q^T Q = I$$

(Q^T também é ortogonal).

E portanto, como se verifica facilmente para qualquer vetor ou matriz x

$$\|Qx\|_2 = \|x\|_2$$

Ou seja, podemos utilizar matrizes ortogonais para transformar o sistema na forma

$$Q^T Ax = Q^T b$$

cujo resíduo

$$r = Q^T b - Q^T Ax$$

tem a mesma norma que o sistema original

$$\|r\|_2 = \|Q^T(b - Ax)\|_2 = \|b - Ax\|_2$$

Ou seja, podemos transformar o sistema em um mais simples de se resolver, com a garantia de que as soluções são a mesma.

Refletores de Householder

Uma transformação ortogonal interessante para o nosso caso é uma reflexão de Householder, que permite a seguinte transformação de um determinado vetor $x \in \mathbb{R}^n$

$$Q \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \rightarrow \begin{bmatrix} -\tau \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

onde Q é o refletor e $\tau = \pm\|x\|_2$. Melhor ainda, sua construção é bem simples, dada por

$$Q = I - \gamma u u^T$$

com

$$\gamma = \frac{x_1 + \tau}{\tau} \text{ , } u = \begin{bmatrix} 1 \\ x_2/(x_1+\tau) \\ \vdots \\ x_n/(x_1+\tau) \end{bmatrix}$$

então só γ e u precisam ser armazenados, e não a Q inteira.[1]

Decomposição QR

Portanto, escolhendo a transformação de forma a realizar isso na primeira coluna da matriz A, com norma τ_1 e a chamando de Q_1, temos

$$Q_1 \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \rightarrow \left[\begin{array}{c|ccc} -\tau_1 & a^*_{12} & \dots & a^*_{1m} \\ \hline 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{array}\right]$$

Podemos ainda escolher uma nova transformação $\hat{Q}_2$ para realizar o mesmo na primeira coluna de A_2, mas precisamos compô-la com a identidade para que ela aja somente em A_2, não alterando o restante de A. Então

$$Q_2 = \left[\begin{array}{c|ccc} I_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \hat{Q}_2 & \\ 0 & & & \end{array}\right]$$

aí teremos

$$Q_2 Q_1 \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \rightarrow \left[\begin{array}{cc|ccc} -\tau_1 & a^*_{12} & a^*_{13} & \dots & a^*_{1m} \\ 0 & -\tau_2 & a^*_{23} & \dots & a^*_{2m} \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & A_3 & \\ 0 & 0 & & & \end{array}\right]$$

Escolhendo uma reflexão para cada coluna de A, poderíamos fazer

$$Q_m \dots Q_1 A = \begin{bmatrix} R \\ 0 \end{bmatrix}$$

com $R \in \mathbb{R}^{m \times m}$ triangular superior. Na prática, no entanto, as colunas de A nem sempre são linearmente independentes, e eventualmente alguma subcoluna terá norma nula (portanto $\tau = 0$), tornando a R singular, o que não permitiria seu uso para resolver o sistema.

[1] Também não é difícil verificar que Q é simétrica e $Q = Q^T$.

Para evitar isso, organizar as colunas de A de forma a manter as subcolunas de maiores normas à esquerda, isto é, antes de escolher cada Q_k $(k = 1 \dots m)$, escolher a subcoluna de maior norma de A_k e trocá-la com a primeira, de forma que os τ_k fiquem organizados em ordem decrescente.

Assim, se no passo $r + 1$ a maior norma das subcolunas de A_{r+1} é zero, então todos os elementos dessa submatriz são nulos e temos

$$Q_r \dots Q_1 AP = R = \begin{bmatrix} R_{11} & R_{22} \\ 0 & 0 \end{bmatrix}$$

onde $R_{11} \in \mathbb{R}^{r \times r}$ é triangular superior e não-singular.[2]

Definindo

$$Q^T = Q_r \dots Q_1$$

temos

$$AP = QR$$

que é a decomposição QR da matriz A com pivotamento.

Solução do Problema de Mínimos Quadrados

Voltando ao sistema original,

$$\begin{aligned} APP^{-1}x &= b \\ QRP^{-1}x &= b \\ RP^{-1}x &= Q^T b \end{aligned}$$

E a resolução do sistema se resume a

$$\begin{aligned} c &= Q^T b \\ R\hat{x} &= c \\ x &= P\hat{x} \end{aligned}$$

Sabendo a forma de R, no entanto, podemos calcular o resíduo [3]

$$\begin{aligned} r &= \begin{bmatrix} R_{11} & R_{22} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} - \begin{bmatrix} \hat{c} \\ d \end{bmatrix} \\ &= \begin{bmatrix} R_{11}\hat{x}_1 + R_{22}\hat{x}_2 - c \\ d \end{bmatrix} \end{aligned}$$

cuja norma ao quadrado é

$$\|r\|_2^2 = \|R_{11}\hat{x}_1 + R_{22}\hat{x}_2 - \hat{c}\|_2^2 + \|d\|_2^2$$

que é quem queremos minimizar. Fica claro que não podemos alterar o termo $\|d\|_2^2$, mas podemos escolher $\hat{x}_1$ e $\hat{x}_2$ de forma a zerar o primeiro termo. Ou seja

$$R_{11}\hat{x}_1 = \hat{c} - R_{22}\hat{x}_2$$

o que tem infinitas soluções, dependendo da escolha dos termos. Como é arbitrário, podemos pegar simplesmente $\hat{x}_2 = 0$ e o segundo passo da solução $(R\hat{x} = c)$ se reduz resolver o sistema triangular superior não-singular

$$R_{11}\hat{x}_1 = \hat{c}$$

e o resíduo se torna

$$\|r\|_2 = \|d\|_2$$

[2] Alterar as colunas de A durante os passos é equivalente a ter multiplicado ela inicialmente por uma matriz de permutação P.

[3] $\hat{x}_1, \hat{c} \in \mathbb{R}^r$

O programa