

Trabalho 1

Prof. Thadeu Dias

2 de novembro de 2023

1 Introdução

1.1 Sistemas sub-determinados

Um problema do tipo $A\vec{x} = \vec{b}$ é dito sub-determinado se o número de incógnitas é maior que o número de equações, $A \in \mathbb{R}^{M \times N}$, $M < N$. Nesse caso, podem haver infinitas soluções, se $\vec{b} \in \mathcal{R}(A)$ ou nenhuma solução no caso contrário. O último é possível se $\text{Rank}(A) \leq M$.

Uma solução do tipo mínimos quadrados *sempre* será solução da equação normal,

$$A^T A \vec{x} = A^T \vec{b}, \quad (1)$$

e se $\vec{b} \in \mathcal{R}(A)$, Um \vec{x} solução da equação acima será uma solução exata. Se $\text{Rank}(A) \leq M$, não existe $(A^T A)^{-1}$, e haverá infinitas soluções possíveis para a equação normal. Se há uma suspeita de que o problema é mal-condicionado ($\text{Rank}(A) \leq M$), um método cuidadoso deve ser usado para se obter uma boa solução para o problema.

1.2 Fatoração QR com permutação de colunas

Assim como a fatoração LU pode sofrer de instabilidades se o elemento na diagonal (o pivô) tiver valor absoluto próximo a zero, a fatoração QR pode se tornar instável se a norma da coluna a ser reduzida for próxima de zero. Uma modificação do algoritmo QR, com permutação de colunas, gera uma matriz triangular superior R , com elementos na diagonal em ordem decrescente de magnitude, e pode ser utilizado com os métodos de Householder ou Givens para se obter uma fatoração QR completa de uma matriz $A \in \mathbb{R}^{M \times N}$ da forma

$$AP = QR, \quad (2)$$

onde $P \in \mathbb{R}^{N \times N}$ é uma matriz de permutação, $Q \in \mathbb{R}^{M \times M}$ é ortogonal, e $R \in \mathbb{R}^{M \times N}$ é triangular superior com elementos na diagonal decrescente em magnitude, isso é,

$$|r_{11}| \geq |r_{22}| \geq |r_{33}| \geq \dots$$

A vantagem da fatoração QR com permutação de colunas é que ela é capaz de revelar o *rank*, da matriz A . Observando os elementos diagonais da matriz R , se A tem rank incompleto, para algum k mínimo, teremos

$$|r_{ii}| \approx 0, \quad i > k.^{[1]}$$

Em outras palavras, o *rank* da matriz A pode ser determinado pelo último elemento da diagonal de R que não se aproxima de zero.

O algoritmo QR com permutação de colunas é uma modificação simples do QR regular. Durante qualquer passo do algoritmo QR, temos

$$Q_m R_m = A P_m. \quad (3)$$

O algoritmo pode ser inicializado com $R_0 = A$, $Q_0 = I_{M \times M}$, $P_0 = I_{N \times N}$. A forma de R é sempre

$$R_m = \left[\begin{array}{c|c} R_{11} & * \\ \hline 0 & R_{22} \end{array} \right], \quad (4)$$

onde $R_{11} \in \mathbb{R}^{m \times m}$ é a região já triangularizada, e $R_{22} \in \mathbb{R}^{M-m \times N-m}$ é a região a ser triangularizada. O algoritmo QR com permutação irá, a cada passo, encontrar uma matriz de permutação P'_{m+1} que trocará a coluna $m+1$ com alguma coluna $j \geq m+1$ de R_m , de forma que a primeira coluna de R'_{22} tenha a maior norma possível, ou seja,

$$Q_m \underbrace{R_m P'_{m+1}}_{R'_m} = A \underbrace{P_m P'_{m+1}}_{P_{m+1}}, \quad (5)$$

onde

$$R'_m = \left[\begin{array}{c|c} R_{11} & * \\ \hline 0 & R'_{22} \end{array} \right].$$

O processo então continua de forma normal, encontrando a matriz ortogonal H_{m+1} que irá aniquilar a $m+1$ -ésima coluna de R'_m

$$\underbrace{Q_m H_{m+1}^T}_{Q_{m+1}} \underbrace{H_{m+1} R'_m}_I = A P_{m+1}.$$

Ao final do processo, temos

$$QR = AP, \quad (6)$$

ou equivalentemente,

$$QRP^T = A,$$

onde R terá a forma

$$R = \left[\begin{array}{c|c} R_{11} & R_{22} \\ \hline 0 & 0 \end{array} \right],$$

^[1]Com precisão infinita, teremos igualdade ao invés de proximidade. Na prática, definimos algum limiar $\epsilon > 0$ de precisão, onde se $|r_{ii}| < \epsilon$, consideramos o valor desprezível.

onde $R_{11} \in \mathbb{R}^{k \times k}$ é um bloco triangular superior com diagonal descendente em magnitude (sendo k o rank de A) e $R_{22} \in \mathbb{R}^{k \times N-k}$ é uma sub-matriz qualquer.

A fatoração QR com permutação pode ser usada como base para se encontrar uma forma prática de se resolver problemas do tipo $A\vec{x} = \vec{b}$ sobre-determinados e com deficiência de rank.

1.3 Decomposição ortogonal completa, UTV^T .

Sabemos que a decomposição QR pode ser usada para resolvermos sistemas do tipo mínimos quadrados de rank-completo: seja $A \in \mathbb{R}^{M \times N}$, $M \geq N$, $\text{Rank}(A) = N$, temos

$$A\vec{x} = \vec{b},$$

fatorando $A = QR$,

$$QR\vec{x} = \vec{b},$$

particionando Q e R ,

$$\begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \vec{x} = \vec{b},$$

reduzindo o sistema

$$\begin{aligned} R_1 \vec{x} &= Q_1^T \vec{b}, \\ \vec{x} &= R_1^{-1} Q_1^T \vec{b}. \end{aligned}$$

onde Q_1 são as primeiras N colunas de Q e R_{11} são as primeiras N linhas de R , e o \vec{x} é a resolução do triangular superior.

Ao mesmo tempo, podemos usar a fatoração QR para sistemas sub-determinados de rank completo, encontrando a solução de norma mínima: seja $A \in \mathbb{R}^{M \times N}$, $M \leq N$, $\text{Rank}(A) = M$

$$A\vec{x} = \vec{b},$$

transpondo,

$$\vec{x}^T A^T = \vec{b}^T,$$

fatorando $A^T = QR$,

$$R^T Q^T \vec{x} = \vec{b},$$

$$Q^T \vec{x} = \vec{y},$$

$$R^T \vec{y} = \vec{b},$$

particionando R e \vec{y} ,

$$\begin{bmatrix} R_{11}^T & 0 \end{bmatrix} \begin{bmatrix} \vec{y}_1 \\ \vec{y}_2 \end{bmatrix} = \vec{b},$$

resolvendo o sistema reduzido,

$$\vec{y} = \begin{bmatrix} R_{11}^{-T} \vec{b} \\ \vec{y}_2 \end{bmatrix}$$

resolvendo para \vec{x} ,

$$\vec{x} = Q \begin{bmatrix} R_{11}^{-T} \vec{b} \\ \vec{y}_2 \end{bmatrix},$$

onde a solução de norma mínima é encontrada exatamente quando $\vec{y}_2 = \vec{0}$.

A idéia da decomposição UTV^T é combinar as duas fatorações, para que possamos resolver sistemas mínimos quadrados de norma mínima.

Seja $A \in \mathbb{R}^{M \times N}$, de rank desconhecido k . Podemos realizar a fatoração QR com permutação em A^T , obtendo

$$\begin{aligned} A^T &= V R P^T, \\ A &= P R^T V^T, \end{aligned}$$

onde $P \in \mathbb{R}^{M \times M}$ é uma matriz de permutação, $V \in \mathbb{R}^{N \times N}$ é uma matriz ortogonal, e R^T terá forma

$$R^T = \left[\begin{array}{c|c} R_{11}^T & 0 \\ \hline R_{21}^T & 0 \end{array} \right],$$

onde R_{11}^T é triangular inferior de tamanho $k \times k$ e R_{21}^T é um bloco qualquer de tamanho $M - k \times k$, ou seja, R^T tem exatamente k colunas linearmente independentes. Podemos então realizar a fatoração QR novamente, em R^T , obtendo

$$A = P R^T V^T,$$

fatorando R^T em U' e T ,

$$A = P U' T V^T,$$

combinando $P U' = U$,^[2]

$$A = U T V^T. \quad (7)$$

Como apenas as primeiras k colunas de R^T são não nulas, T será da forma

$$T = \left[\begin{array}{c|c} T_{11} & 0 \\ \hline 0 & 0 \end{array} \right], \quad (8)$$

onde T_{11} é o bloco triangular superior de tamanho $k \times k$.

1.3.1 Solução mínimos quadrados para sistemas de rank incompleto

A solução para sistemas inconsistentes subdeterminados é exatamente a combinação dos dois métodos QR. Seja $A \in \mathbb{R}^{M \times N}$ de rank incompleto. Resolver um sistema do tipo $A\vec{x} = \vec{b}$ consiste nos seguintes passos:

$$A\vec{x} = \vec{b},$$

^[2]Observe que P também é ortogonal.

fatorando A ,

$$UTV^T \vec{x} = \vec{b},$$

seja $V^T \vec{x} = \vec{y}$,

$$UT\vec{y} = \vec{b},$$

particionando

$$\left[\begin{array}{c|c} U_1 & U_2 \end{array} \right] \left[\begin{array}{c|c} T_{11} & 0 \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c} \vec{y}_1 \\ \vec{y}_2 \end{array} \right] = \vec{b},$$

reduzindo,

$$U_1 \left[\begin{array}{c|c} T_{11} & 0 \end{array} \right] \left[\begin{array}{c} \vec{y}_1 \\ \vec{y}_2 \end{array} \right] = \vec{b},$$

projetando,

$$\left[\begin{array}{c|c} T_{11} & 0 \end{array} \right] \left[\begin{array}{c} \vec{y}_1 \\ \vec{y}_2 \end{array} \right] = U_1^T \vec{b},$$

solução de norma mínima,

$$\vec{y} = \left[\begin{array}{c} T_{11}^{-1} U_1^T \vec{b} \\ 0 \end{array} \right],$$

resolvendo para \vec{x} ,

$$\vec{x} = V \left[\begin{array}{c} T_{11}^{-1} U_1^T \vec{b} \\ 0 \end{array} \right].$$

O vetor \vec{x} encontrado dessa forma é exatamente a solução do problema de otimização com restrições

$$\begin{aligned} & \arg \min_{\vec{x}} \quad \|\vec{x}\| \\ & \text{sujeito a} \quad A^T A \vec{x} = A^T \vec{b} \end{aligned} \tag{9}$$

Além disso, considere a matriz

$$\begin{aligned} A^\dagger &= V \left[\begin{array}{c|c} T_{11}^{-1} & 0 \\ \hline 0 & 0 \end{array} \right] U^T \\ &= VT^\dagger U^T. \end{aligned}$$

Observe que

$$\begin{aligned} AA^\dagger A &= (UTV^T)(VT^\dagger U^T)(UTV^T) \\ &= UTT^\dagger TV^T \\ &= U \left[\begin{array}{c|c} T_{11} & 0 \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c|c} T_{11}^{-1} & 0 \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c|c} T_{11} & 0 \\ \hline 0 & 0 \end{array} \right] V^T \\ &= U \left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c|c} T_{11} & 0 \\ \hline 0 & 0 \end{array} \right] V^T \\ &= UTV^T \\ &= A. \end{aligned}$$

De forma similar, pode ser demonstrado que

$$\begin{aligned} A^\dagger A A^\dagger &= A^\dagger, \\ (A A^\dagger)^T &= A A^\dagger, \\ (A^\dagger A)^T &= A^\dagger A, \end{aligned}$$

ou seja, A^\dagger é a pseudo-inversa de A .

2 Elaboração do trabalho

Realize as seguintes tarefas, e faça um relatório descrevendo o trabalho realizado:

- Implemente a fatoração QR sem permutação, usando o método de Householder ou Givens. Verifique se para seu método, $A = QR$ com diferentes tamanhos de matrizes.
- Implemente a fatoração QR com permutação, novamente, usando Householder ou Givens. Verifique se $AP = QR$, e se a diagonal de R é decrescente em módulo.
- Implemente usando os dois métodos do item anterior, a fatoração UTV^T . Novamente, verifique se $UTV^T = A$, para diferentes tamanhos de matrizes. Verifique se T tem a forma triangular em blocos como esperado.
- Implemente o algoritmo de back-substitution (triangular superior), e use o método descrito acima para resolver problemas de 4 tipos:
 1. $A\vec{x} = \vec{b}$, A é quadrada de rank completo.
 2. $A\vec{x} = \vec{b}$, A é retangular alta e estreita de rank completo.
 3. $A\vec{x} = \vec{b}$, A é retangular baixa e larga de rank completo.
 4. $A\vec{x} = \vec{b}$, A é quadrada de rank incompleto.
 5. $A\vec{x} = \vec{b}$, A é retangular, baixa e larga de rank incompleto.

Para gerar matrizes de tamanho $M \times N$, de rank k , use

$$A = \sum_{i=1}^k \vec{u}_i \vec{v}_i^T,$$

onde \vec{u}_i e \vec{v}_i são vetores aleatórios de tamanho M e N respectivamente. O vetor \vec{b} deve ser gerado aleatoriamente. Descreva, em cada caso, o comportamento da solução \vec{x} : $A\vec{x}$ é próximo de \vec{b} ? Qual a norma de \vec{x} ?