Solução do Problema de Quadrados Mínimos Discretos Utilizando a Decomposição em Valores Singulares (SVD)

João Pedro de Lima e Raymundo Eduardo Pilz

2025-06-23

1. Introdução

O método dos mínimos quadrados é uma das ferramentas mais fundamentais da estatística e da análise numérica, sendo utilizado para ajustar modelos a dados experimentais. Sua principal aplicação consiste em encontrar uma função que melhor represente um conjunto de observações, minimizando os erros entre os valores observados e os valores previstos pelo modelo.

Historicamente, o método foi desenvolvido de forma independente por Adrien-Marie Legendre e Carl Friedrich Gauss. Este último, em particular, aplicou a técnica para prever a órbita de corpos celestes, como o asteroide Ceres, a partir de observações incompletas.

O problema de mínimos quadrados pode ser formulado, em termos matriciais, como a busca por um vetor de parâmetros $\mathbf{x}$ que minimize a norma euclidiana do resíduo:

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$$

onde $\mathbf{A}$ é a matriz de design e $\mathbf{b}$ é o vetor de respostas observadas.

Este trabalho tem como foco a resolução desse problema utilizando a Decomposição em Valores Singulares (SVD), destacando suas vantagens numéricas e computacionais em relação a outras abordagens clássicas.

2. Fundamentação Teórica

2.1 O Problema de Quadrados Mínimos Discretos

Dado um conjunto de observações experimentais representadas por pares ordenados (x_i, y_i), com $i = 1, \dots, n$, o objetivo é encontrar os parâmetros de um modelo $f(x, \beta)$ que melhor se ajustem aos dados. O critério de ajuste é baseado na minimização da soma dos quadrados dos resíduos:

$$S(\beta) = \sum_{i=1}^{n} [y_i - f(x_i, \beta)]^2$$

Em casos simples, como a regressão linear, o modelo assume a forma:

$$f(x, \beta) = \beta_0 + \beta_1 x$$

2.2 Alternativas para a Resolução

Existem diferentes métodos para resolver o problema de mínimos quadrados:

- **Equações Normais:**
  $$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$
  Método rápido, porém sensível a problemas de condicionamento numérico.
- **Decomposição QR:** Boa estabilidade numérica, mas menos robusta que a SVD frente a colinearidade severa.
- **Decomposição SVD:** Extremamente estável numericamente. Recomendada para problemas com colinearidade, dados mal-condicionados ou de posto deficiente.

Além desses, métodos iterativos e regularizados (como Ridge Regression) podem ser utilizados em casos específicos.

2.3 A Decomposição SVD

A Decomposição em Valores Singulares expressa qualquer matriz $A \in \mathbb{R}^{m \times n}$ como o produto:

$$A = U \Sigma V^T$$

Onde:

- U é uma matriz ortogonal de dimensão $m \times m$,
- Σ é uma matriz diagonal $m \times n$ contendo os valores singulares,
- V é uma matriz ortogonal de dimensão $n \times n$.

Entre as principais propriedades da SVD, destacam-se:

- Robustez numérica;
- Capacidade de identificar direções de variabilidade dos dados;
- Permite o cálculo da pseudo-inversa de maneira estável.

2.4 Condicionamento Numérico e Estabilidade da SVD

O condicionamento numérico de um problema está relacionado à sensibilidade de sua solução em relação a pequenas perturbações nos dados de entrada. No contexto de sistemas lineares e mínimos quadrados, o número de condição da matriz $\mathbf{A}$ desempenha papel central.

O número de condição $\kappa(\mathbf{A})$ na norma 2 é definido como:

$$\kappa(\mathbf{A}) = \frac{\sigma_1}{\sigma_r}$$

onde σ_1 é o maior valor singular e σ_r é o menor valor singular não nulo de $\mathbf{A}$.

Um número de condição alto (muito maior que 1) indica que o problema é mal condicionado, ou seja, pequenas alterações em $\mathbf{b}$ podem causar grandes mudanças na solução $\mathbf{x}$. Isso é comum em problemas com colinearidade entre variáveis ou dados mal escalados.

A SVD permite identificar e lidar com o mal condicionamento, pois:

- Os valores singulares mostram diretamente a escala de variabilidade das direções no espaço de colunas de $\mathbf{A}$;
- Componentes com σ_i muito pequenos podem ser eliminados ou regularizados (por exemplo, com truncamento da SVD ou regressão ridge), melhorando a estabilidade da solução;
- Evita a inversão direta de $\mathbf{A}^T\mathbf{A}$, que pode amplificar o erro se $\mathbf{A}$ for mal condicionada.

Dessa forma, a decomposição SVD é especialmente adequada para resolver o problema de mínimos quadrados com robustez numérica, mesmo quando a matriz $\mathbf{A}$ é de posto deficiente ou próxima de singular.

3. Metodologia

SVD é uma técnica matricial que decompõe uma matriz (real ou complexa) original em três componentes:

- Uma matriz ortogonal U
- Uma matriz diagonal Σ
- A transposta da matriz ortogonal V

A representação geral da SVD é dada por:

$$A = U\Sigma V^T$$

Afim de contextualizar a base da construção da SVD, tem-se o **Teorema SVD Geométrico:**

Seja $A \in \mathbb{R}^{n\times m}$ uma matriz de posto $r > 0$. Então, existem $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$, uma base ortonormal $\{v_1, ..., v_m\}$ de $\mathbb{R}^m$ e uma base ortonormal $\{u_1, ..., u_n\}$ de $\mathbb{R}^n$ de modo que

$$Av_i = \begin{cases} \sigma_i u_i, & \text{se } i = 1, ..., r \\ 0, & \text{se } i = r+1, ..., m. \end{cases} \quad \text{e} \quad A^T u_i = \begin{cases} \sigma_i v_i, & \text{se } i = 1, ..., r \\ 0, & \text{se } i = r+1, ..., n. \end{cases}$$

Em particular, $v_1, ..., v_m$ são autovetores de A^TA, $u_1, ..., u_n$ são autovetores de AA^T, e $\sigma_1^2, ..., \sigma_r^2$ são os autovalores não nulos de A^TA e AA^T.

Além disso:

Para cada $A \in \mathbb{R}^{n\times m}$ de posto r, existe uma matriz

$$\Sigma = \begin{pmatrix} \hat{\Sigma} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n\times m},$$

em que $\hat{\Sigma} = \text{diag}\{\sigma_1, ..., \sigma_r\}$ e $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$, e existem matrizes ortogonais $U \in \mathbb{R}^{n\times n}$ e $V \in \mathbb{R}^{m\times m}$ tais que

$$A = U\Sigma V^T.$$

Para cada $A \in \mathbb{R}^{n\times m}$ de posto r, existe uma matriz

$$\Sigma = \begin{pmatrix} \hat{\Sigma} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n\times m},$$

em que $\hat{\Sigma} = \text{diag}\{\sigma_1, ..., \sigma_r\}$ e $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$, e existem matrizes ortogonais $U \in \mathbb{R}^{n\times n}$ e $V \in \mathbb{R}^{m\times m}$ tais que

$$A = U\Sigma V^T.$$

Os números $\sigma_1, \sigma_2, ..., \sigma_r$ são denominados de valores singulares de A. As colunas de U são chamadas de vetores singulares à esquerda e as colunas de V de vetores singulares à direita. O produto $U\Sigma V^T$ é denominado uma **SVD** de A.

Cada uma das matrizes envolvidas possui propriedades específicas que desempenham um papel crucial na decomposição:

Matriz U:

- As colunas de U são autovetores de AA^T.
- $\det(U) = \pm 1$
- $U^{-1} = U^T$
- U forma uma base ortogonal para o espaço das colunas de A.

Matriz Σ:

- Os valores singulares σ_i são raízes quadradas dos autovalores de $A^T A$.
- A matriz Σ é diagonal, o que significa que todos os seus elementos fora da diagonal principal são zero.
- Os elementos da diagonal principal de Σ, denotados como $\sigma_1, \sigma_2, \ldots, \sigma_r$, são os valores singulares de A, onde r é o posto da matriz A. Esses valores são dispostos em ordem decrescente: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq 0$.
- A matriz Σ terá $\min(m, n)$ valores singulares, onde os valores além do posto da matriz (se houver) são zero.

Se A é uma matriz $m \times n$:

- Σ terá m linhas e n colunas.
- Se $m > n$, Σ terá n valores singulares seguidos de $m - n$ zeros.
- Se $m < n$, Σ terá m valores singulares seguidos de $n - m$ zeros.

Matriz V:

- $\det(V) = \pm 1$
- $V^{-1} = V^T$
- V forma uma base ortogonal para o espaço das linhas de A.
- A matriz V é construída a partir de um conjunto ortonormal de autovetores da matriz $A^T A$.

As matrizes U e V são ortogonais, ou seja:

$$A^{-1} = A^T \iff AA^T = I$$

- Para a construção de Σ na fatoração, temos que encontrar os autovalores da matriz $(A^T A)$, ordenamos esses autovalores do maior para o menor, denotamos como:

$\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_k^2 > \sigma_{k+1} = \cdots = \sigma_n = 0$

onde σ_k^2 é o menor autovalor maior que 0 de $A^T A$.

A matriz diagonal Σ é dada por:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

onde $\sigma_i = 0$ quando $k < i \leq n$. Os valores singulares são únicos para cada matriz A.

- Para a construção da matriz ortogonal V, usamos a matriz simétrica $A^T A$, fatoramos ela como:

$$A^T A = V D V^T$$

V é uma matriz ortogonal $n \times n$. Cujas colunas são chamadas de vetores singulares direitos, as colunas são os autovetores de $A^T A$. Os vetores singulares direitos formam uma base ortonormal para o espaço linha de A.

- Os elementos de U são determinados por

$$u_i = \frac{1}{\sigma_i} A v_i, \quad i = 1, 2, \dots, n$$

U é uma matriz $m \times m$, onde as colunas são chamadas de vetores singulares esquerdos, os vetores singulares esquerdos representam uma base ortonormal para o espaço da coluna de A.

Determinamos os autovetores de AA^T, as colunas desses autovetores formam a matriz U. É comum selecionar os autovetores correspondentes aos maiores autovalores para as primeiras colunas de U. Para termos a garantia de que U é ortogonal nós normalizamos cada coluna de U dividindo pela sua norma Euclidiana.

Deste modo, a matriz A $m \times n$, com n valores singulares pode ser escrita na forma:

$$A = \sigma_1 u_1 v_1^T + \dots + \sigma_n u_n v_n^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

Pode-se implementar um **algoritmo** para a solução de norma mínima para o problema de mínimos quadrados via SVD a partir de:

- Calcular a SVD de A;
- Calcular o vetor $c = U^T b$;

- Calcular o vetor $y = (c_1/\sigma_1, \dots, c_m/\sigma_m)$ (se $r = m$) e $y = (c_1/\sigma_1, \dots, c_r/\sigma_r, 0, \dots, 0)$ (se $r < m$);
- Calcular o vetor $x = Vy$.

O problema de Mínimos Quadrados via SVD

- $Ax = b$ nem sempre tem solução exata.
- O problema de mínimos quadrados consiste em determinar um vetor x que minimiza a expressão $\|b - Ax\|_2$.
- A denominação desse problema se deve ao fato de estarmos minimizando a soma dos quadrados dos resíduos.

A partir isso, temos o **Teorema** a seguir:

Sejam $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$ $(m > n)$ e posto$(A) = r$. O problema de mínimos quadrados associado ao sistema linear $Ax = b$ possui exatamente uma solução de norma mínima.

A resolução do problema de mínimos quadrados via SVD segue os seguintes passos:

- Cálculo da SVD de A:
$$A = U\Sigma V^T$$
- Transformação do vetor de respostas:
$$\mathbf{c} = U^T\mathbf{b}$$
- Cálculo do vetor intermediário $\mathbf{y}$: Se o posto de A for r:
$$y_i = \frac{c_i}{\sigma_i}, \quad i = 1, \dots, r$$
- Determinação da solução:
$$\mathbf{x} = V\mathbf{y}$$

  Esse processo resulta na soluçãoo de norma mínima para o problema.

4. Resultados

A resolução de um problema de regressão linear simples pelo método dos mínimos quadrados via SVD, envolve a construção da matriz de design, a aplicação da decomposição SVD, o cálculo da pseudo-inversa e, finalmente, a multiplicação pela matriz de resultados.

A metodologia foi implementada em linguagem **Julia** para os dados de consumo e renda, segue script abaixo:

4.1 Exemplo 1: Relação entre Consumo e Renda

Utilizando dados fictícios de consumo ($\mathbf{y}$) e renda ($\mathbf{x}$):

i	**Consumo** (y)	**Renda** (x)
1	122	139
2	114	126
3	86	90
4	134	144
5	146	163
6	107	136
7	68	61
8	117	62
9	71	41
10	98	120

Tabela 1: Dados de Consumo e Renda, origem: Wikipedia

Implementação em Julia:

```julia
## Dados da Tabela
x = [139.0, 126.0, 90.0, 144.0, 163.0, 136.0, 61.0, 62.0, 41.0, 120.0]
y = [122.0, 114.0, 86.0, 134.0, 146.0, 107.0, 68.0, 117.0, 71.0, 98.0]

## Número de observações
n = length(x)

## Construir a Matriz de Design A e o Vetor b
A = hcat(ones(n), x)
b = y

## Realiza a SVD de A
F = svd(A)

## Constrói a matriz Sigma_plus (pseudo-inversa dos valores singulares)
tolerance = max(size(A)...) * eps(Float64) * F.S[1]
Sigma_plus_diag = [s > tolerance ? 1.0/s : 0.0 for s in F.S]

## A pseudo-inversa A_plus
A_plus = F.V * Diagonal(Sigma_plus_diag) * F.U`

## Encontrar os Coeficientes Beta
beta_hat = A_plus * b

println("Coeficientes estimados (beta_0, beta_1):")
println("beta_0 (intercepto) = ", beta_hat[1])
```

```
println("beta_1 (coeficiente de Renda) = ", beta_hat[2])
```

Como mencionado anteriormente, a solução do problema habilita diversas possibilidades para tratativa, há o exemplo de **Previsão**, que pode ser implementado da seguinte maneira:

```
## Função para prever consumo
function predict_consumption(renda)
    return beta_hat[1] + beta_hat[2] * renda
end

## Coeficientes estimados (beta_0, beta_1):
beta_0 (intercepto) = 52.69018446221764
beta_1 (coeficiente de Renda) = 0.4954054002447952

predicted_y = predict_consumption(100.0)
println("Consumo previsto para Renda = 100: ", predicted_y)

## Estimativa para a população do Brasil em 2025 (exemplo anterior)
println("Estimativa em 2025: ", predict_consumption(2025.0))

## Consumo previsto para Renda = 100: 102.23072448669716
## Estimativa em 2025: 1061.2185590924962
```

Coeficientes estimados: beta_0 (intercepto) = 52.69018446221764 beta_1 (coeficiente de Renda) = 0.4954054002447952

Resultado:

$$\hat{\beta}_0 \approx 52.69, \quad \hat{\beta}_1 \approx 0.4954$$

4.2 Exemplo 2: Crescimento Populacional com dados de população ao longo dos anos.

Modelando um crescimento exponencial:

$$P(t) = ae^{bt}$$

Após transformação logarítmica e aplicação da SVD:

```
anos = [2000, 2005, 2010, 2015, 2020]
pop = [1000, 1200, 1450, 1700, 2000]

t = anos .- minimum(anos) # Transformar anos em tempo relativo
y_log = log.(pop) # Transformar população para log
```

```
A_exp = hcat(ones(length(t)), t) # Matriz de design para o modelo linearizado
U, S, Vt = svd(A_exp)

tolerance_exp = max(size(A_exp)...) * eps(Float64) * S[1]
Sigma_plus_exp_diag = [s > tolerance_exp ? 1.0/s : 0.0 for s in S]
A_plus_exp = Vt * Diagonal(Sigma_plus_exp_diag) * U'

x_coeffs = A_plus_exp * y_log
c0, c1 = x_coeffs

a_exp = exp(c0) # Voltar para a escala original para 'a'
b_exp = c1      # 'b' permanece o mesmo

println("Coeficientes do modelo exponencial:")
println("a = ", a_exp)
println("b = ", b_exp)
```

Coeficientes do modelo exponencial: a = 998.6749007624955 b = 0.0354143423719998

5.Conclusão

O problema dos mínimos quadrados discretos, ao buscar minimizar a soma dos quadrados dos resíduos, apresenta diferentes estratégias de solução. Dentre elas, a Decomposição em Valores Singulares (SVD) destaca-se pela robustez numérica, tornando-se especialmente vantajosa em situações de colinearidade ou dados mal-condicionados.

Através de exemplos práticos, foi possível demonstrar como a SVD permite obter soluções estáveis e interpretar geometricamente a estrutura dos dados. O uso da linguagem Julia, com seu suporte eficiente a operações matriciais, também contribuiu para a aplicabilidade computacional dos métodos discutidos.

Como possiveis desdobramentos, recomenda-se explorar abordagens baseadas em regularização, como Ridge Regression ou métodos iterativos, para comparação dos métodos, e a aplicação para resolução de problemas específicos e definidos, por exemplo, otimização de carteiras de seguros, utilizando otimização estocástica.

Dentre as dificuldades encontradas durante o desenvolvimento do trabalho, houve a dificuldade de compreensão de conceitos matemáticos, até então novos que exigiram estudos de diferente maneiras, sejam por leitura, vídeo-aulas, e a busca para encontrar bibliografia que apresentasse definições formais. Além disso, a formatação desse documento se mostrou desafiadora, até então, era uma maneira desconhecida, mas se apresentou muito prática utilizando recursos do LaTeX via documentos qmd, digitados na IDE RStudio

6. Referências bibliográficas

STRANG, Gilbert. Álgebra Linear e suas Aplicações. Tradução da 4ª ed. norte-americana. São Paulo: Cengage Learning, 2010.

DECOMPOSIÇÃO EM VALORES SINGULARES. In: WIKIPÉDIA: a enciclopédia livre. [S.l.]: Wikimedia Foundation, [s.d.]. Disponível em: \url{https://pt.wikipedia.org/wiki/Decomposi%C3%A7%C3%A3o_em_Valores_Singulares}. Acesso em: 23 jun. 2025.

MÉTODO DOS MÍNIMOS QUADRADOS. In: WIKIPÉDIA: a enciclopédia livre. [S.l.]: Wikimedia Foundation, [s.d.]. Disponível em: \url{https://pt.wikipedia.org/wiki/M%C3%A9todo_dos_m%C3%ADnimos_quadrados}. Acesso em: 23 jun. 2025.