Ajuste Automático ao Histórico
Simulação de Reservatórios

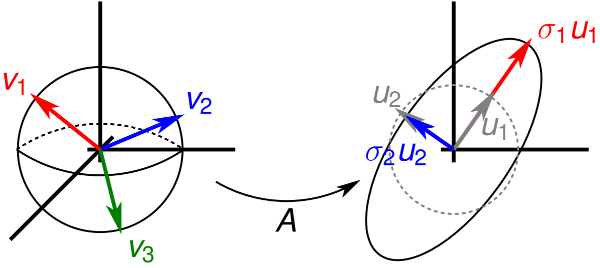
Flavio Dickstein IM/UFRJ
Paulo Goldfeld IM/UFRJ
Renan V. Pinto LabMAPetro/UFRJ

Novembro/2018

Plano

SVD

- Decomp. espectral: diagon. de matriz **quadrada.**
  $A = MDM^{-1}, \quad M$ é mudança de base.
  Nem toda matriz quadrada é diagonalizável.
- SVD: diagonalização de **uma matriz qualquer**.
  Mudança**s** de base **ortogonais**.



SVD

$A = U\Sigma V^T$

- $A_{m\times n}$ é qualquer;
- $U_{m\times m}$ é ortogonal, $U^T U = I$;
  $V_{n\times n}$ é ortogonal, $V^T V = I$;
- $\Sigma_{m\times n}$ diagonal, $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$.

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \\ & & \end{bmatrix} \text{ ou } \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_m & & \end{bmatrix}.$$

$A = U\Sigma V^T \Leftrightarrow AV = U\Sigma \Leftrightarrow Av_i = \sigma_i u_i, i \leq n$

u_i's e v_i's são os vetores singulares à esquerda e à direita, σ_i são os valores singulares.

Aproximação de A

Seja $k < n$

$$\hat{\Sigma}_k = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_k & \\ \hline & 0 & & 0 \end{array}\right] = \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix}$$

$$A_k = U\hat{\Sigma}_k V^T = \begin{bmatrix} U_k & \tilde{U} \end{bmatrix} \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_k^T \\ \tilde{V}^T \end{bmatrix} = U_k \Sigma_k V_k^T$$

Se $\sigma_k > 0$, A_k é uma aproximação de A de posto k.

Aproximação de A

A_k é a matriz de posto k "mais próxima de A". Defina

$$\|u\|_2 = \left(\sum_i |u_i|^2\right)^{1/2}$$

$$\|A\|_2 = \max\{\|Au\|_2, \|u\|_2 = 1\} \quad \text{(norma 2)}$$

$$\|A\|_F = \left(\sum_{i,j} A_{i,j}^2\right)^{1/2} \qquad \text{(norma de Frobenius)}$$

Então, se $\operatorname{rank}(B) \leqslant k$,

$\|A - B\|_2 \geqslant \|A - A_k\|_2 = \sigma_{k+1}$.

Se $B \neq A_k$, $\|A - B\|_F > \|A - A_k\|_F = \left(\sum_{i>k} \sigma_i^2\right)^{1/2}$

Os valores e espaços singulares são únicos.

Número de condicionamento

Considere os sistemas lineares invertíveis

$$Ax = b \quad \text{e} \quad A(x + \delta x) = b + \delta b$$

de modo que

$$A\delta x = \delta b$$

Defina os erros relativos

$$e_b = \frac{\|\delta b\|}{\|b\|}, \quad e_x = \frac{\|\delta x\|}{\|x\|}$$

Então,

$$\frac{e_x}{e_b} = \frac{\|\delta x\|}{\|\delta b\|} \frac{\|b\|}{\|x\|} = \frac{\|A^{-1}\delta b\|}{\|\delta b\|} \frac{\|Ax\|}{\|x\|}$$

de modo que

$$\max_{x,\delta b} \frac{e_x}{e_b} = \|A\| \|A^{-1}\|$$

Número de condicionamento

Definimos o **número de condicionamento** $\kappa(A)$ como

$$\kappa(A) = \|A\|\|A^{-1}\|.$$

O condicionamento de A é o fator de amplificação do erro relativo introduzido na resolução de um sistema.

$\kappa(A)$ depende da norma considerada. No caso da norma 2,

$$\kappa_2(A) = \|A\|_2\|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}.$$

é o quociente entre o maior e o menor valor singular de A.

Número de condicionamento

Quando A é singular, consideramos o problema:
"encontrar o 'menor' x tal que $\|b - Ax\|$ seja mínima".

Neste caso, se $\text{rank}(A) = k$, temos

$$\kappa_2(A) = \frac{\sigma_1}{\sigma_k}.$$

Se A tem posto n e aproximamos A por A_k, sua melhor representação de posto $k < n$, melhoramos o condicionamento do problema.

LabMAPetro

Métodos com derivada

Ideia básica: f_k aproximação quadrática de f em torno de m_k.

$$f_k(m_k + \delta m) = f(m_k) + \nabla f(m_k)^T \delta m + \frac{1}{2}\delta m^T \nabla^2 f(m_k)\delta m$$

Minimizando f_k:

$$\nabla f_k(m_k + \delta m_k)^T = 0 = \nabla f(m_k)^T + \nabla^2 f(m_k)\delta m_k.$$

Assim, $\delta m_k = -\left(\nabla^2 f(m_k)\right)^{-1} \nabla f(m_k)^T$.

Newton:	$+$	convergência quadrática;
	$-$	não garante convergência (nem mesmo descida) e
	$-$	exige o cálculo de $\nabla^2 f(m_k)$.

Otimização Não-Linear

Método de Newton: $\delta m_k = -\left(\nabla^2 f(m_k)\right)^{-1} \nabla f(m_k)^T$

Métodos “tipo” Newton: $\delta m_k = -\alpha H_k^{-1} \nabla f(m_k)^T.$

- $H_k > 0 \;\Rightarrow\;$ direção de descida;
- escolha de $\alpha > 0$ é a busca linear;
- Newton: $\alpha = 1,\; H_k = \nabla^2 f(m_k)$;
- Máxima Descida: $H_k = I$ e
- Quasi-Newton: $H_k > 0$ aproxima $\nabla^2 f(m_k)$, usando $\nabla f(m_j),\; j \leq k$.

Método de Gauss-Newton

O Método de Gauss-Newton se aplica a problemas de mínimos quadrados não-lineares

$$\min_{m} \frac{1}{2}\|r(m)\|^2$$

A cada iteração k aproximamos

$$r(m_k + \delta m) \approx r(m_k) + r'(m_k)\delta m = A\,\delta m - b$$

e resolvemos o problema de mínimos quadrados **linear**

$$\min_{\delta m} \frac{1}{2}\|A\,\delta m - b\|^2$$

cuja solução é dada pelas **equações normais**

$$A^T A\,\delta m_k = A^T b$$

Probabilidade

Um espaço de medida (E, Σ, P) é dito um espaço de probabilidade se $P \geqslant 0$ e $P(E) = 1$.

Uma variável aleatória (v.a.) é uma função mensurável $X : E \to \mathbb{R}^n$.

No que segue $E = \mathbb{R}^m$, Σ é a σ-álgebra de Borel e μ é a medida de Lebesgue.

Seja X uma v.a. Então, $f_X : \mathbb{R} \to \mathbb{R}$ é uma densidade de probabilidade (dp) de X se

$$P(X(x) \in \Omega) = \int_\Omega f_X(x)\, dx.$$

Esperança e Variância

A **média** μ (ou **valor esperado** $E(X)$) de X é dada por

$$\mu = \int_{\mathbb{R}^m} x f_X(x)\, dx \in \mathbb{R}^n$$

Se X é escalar, a **variância** σ^2 de X é dada por

$$\sigma^2 = E((X-\mu)^2) = \int_{\mathbb{R}^m} (x-\mu)^2 f_X(x)\, dx$$

Se $X = (X_1, X_2, \ldots, X_n)$, a **matriz de covariância** C de X é

$$C_{i,j} = E((X_i - \mu_i)(X_j - \mu_j))$$

Seja $X = (X_1, X_2)$. Então, X_1, X_2 são v.a. **independentes** se $f_X(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$, isto é,

$$P(X_1 \in \Omega_1 \wedge X_2 \in \Omega_2) = P(X_1 \in \Omega_1) P(X_2 \in \Omega_2)$$

Neste caso, $C(X)$ é diagonal.

Distribuição Normal

Se $\sigma \neq 0$, a densidade normal $\mathcal{N}_{\mu,\sigma^2}$ dada por

$$f(x) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{\langle x-\mu, x-\mu \rangle}{2\sigma^2}} \propto e^{-\frac{\|x-\mu\|^2}{2\sigma^2}}$$

tem média μ e variância σ^2. $\mathcal{N}_{0,1}$ é dita normal padrão.

Se $C_{n\times n}$ é SPD, então $\mathcal{N}_{\mu,C}$ dada por

$$f(x) \propto e^{-\frac{1}{2}\langle C^{-1}(x-\mu), x-\mu \rangle}$$

tem média $\mu \in \mathbb{R}^n$ e covariância C. $\mathcal{N}_{0,I}$ é dita normal padrão.

Se X tem distribuição $\mathcal{N}_{\mu,C}$, então X_i, $i \leq n$, são independentes se e só se C é diagonal.

Distribuição χ^2

Seja $C = C^{1/2}C^{1/2}$, $C^{1/2}$ SPD. Então,
$\langle C^{-1}(x-\mu), x-\mu\rangle = \|C^{-1/2}(x-\mu)\|^2$.
Se X é uma v.a. de densidade $\mathcal{N}_{\mu,C}$, então $Y = C^{-1/2}(X-\mu)$ é normal padrão.

Sejam $X_1, \ldots, X_n$ v.a. independentes com distribuição normal padrão. Então,

$$\|X\|^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

é dita uma v.a. χ_n^2. $\|X\|^2$ tem média n e variância $\sigma^2 = 2n$.

Se X é $\mathcal{N}_{\mu,C}$, então,

$$\|C^{-1/2}(X-\mu)\|^2$$

tem distribuição χ_n^2.