矩阵论 第七讲

上次：奇异值分解（第五章）

5.2.2 酉不变性（奇异值）

$$\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^H$$

$$\underline{\underline{B}} = \underline{\underline{P}}\,\underline{\underline{A}}\,\underline{\underline{Q}}^H = \underline{\underline{P}}\,\underline{\underline{U}}\,\underline{\underline{\Sigma}}(\underline{\underline{Q}}\,\underline{\underline{V}})^H$$

酉矩阵

奇异值不变：酉不变性

奇异向量不同

5.2.3 秩亏缺 LS（不是满秩） 应用奇异值分解求解最小二乘问题

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

第一种方法：$\underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^H \underline{x} = \underline{b}$ $\underline{\underline{\Sigma}}\,\underline{\underline{V}}^H \underline{x} = \underline{\underline{U}}^H \underline{b}$ 令 $\underline{x}' = \underline{\underline{V}}^H \underline{x}$, $\underline{y} = \underline{\underline{U}}^H \underline{b}$

$$\| \underline{\underline{\Sigma}}\,\underline{x}' - \underline{\underline{U}}^H \underline{b} \|_2^2 \to \min$$

$$\sum_{i=1}^{r} |\sigma_i x_i' - y_i|^2 + \sum_{i=r+1}^{N} |y_i|^2$$

$\sum_{i=1}^{r} |\sigma_i x_i' - y_i|^2 = 0$

求出 $x_i' \to x'$

第二种方法：

$$\underline{\underline{A}}\,\underline{x} = \underline{b} \Rightarrow \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^H \underline{x} = \underline{b}$$

$$\underline{x} = \underline{\underline{G}}\,\underline{b} = (\underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^H)^{\dagger} \underline{b} = \underline{\underline{V}}\,\underline{\underline{\Sigma}}^{\dagger}\,\underline{\underline{U}}^H \underline{b}$$

$$\underline{\underline{\Sigma}}^{\dagger} = \mathrm{diag}(1/\sigma_1, 1/\sigma_2 \cdots 1/\sigma_r, 0 \cdots 0)$$

$$\underline{x} = \sum_{i=1}^{r} \left(\frac{\underline{u}_i^H \underline{b}}{\sigma_i}\right) \underline{v}_i$$ 最小二乘问题的最小范数解

相应的最小残差为：

$$\rho_{LS} = \| [\underline{u}_{r+1}, \underline{u}_{r+2}, \cdots \underline{u}_m]^H \underline{b} \|_2$$

$$\underline{\underline{A}} = \sum_{i=1}^{r} \sigma_i \underline{u}_i \underline{v}_i^H$$

矩阵近似：

$\underline{\underline{B}}$ 矩阵近似 $\underline{\underline{A}}$.

如何衡量是否准.

$\mathrm{rank}(\underline{\underline{B}}) = K$. 先确定 B 的秩

定理：$\underline{A} \in R^{m\times n}$ 的奇异值分解由 $\underline{A} = \sum_{i=1}^{p} \sigma_i \underline{u}_i \underline{v}_i^H$ 给出，其中 $p = \mathrm{rank}(\underline{A})$. 若 $k<p$. 且 $\underline{B} = \sum_{i=1}^{k} \sigma_i \underline{u}_i \underline{v}_i^H$. 则逼近质量分别使用谱范数和 Frobenius 范数衡量

$$\min_{\mathrm{rank}(\underline{B})=k} \|\underline{A}-\underline{B}\|_{spec} = \sigma_{k+1}.$$

$$\min_{\mathrm{rank}(\underline{B})=k} \|\underline{A}-\underline{B}\|_F = \sqrt{\sum_{i=k+1}^{q} \sigma_i^2} \qquad q = \min\{m,n\}$$

理论上 $i>r$ 时. $\sigma_i = 0$.

实际上 $\sigma_i \neq 0$. 有时甚至出现比较大的扰动

估计有效秩.

确定有两种方法：

有效秩确定

$$\varepsilon = \frac{\sigma_k}{\sigma_1}$$ 归一化奇异值方法 最大整数 $\varepsilon \geq 5\%$ 判断在哪里截断

$$\underline{B} = \underline{U}_k \underline{\Sigma}_k \underline{V}_k^H$$

令m×n矩阵 A_k 是原m×n矩阵A的秩k近似

范数比方法:

$$\nu(k) = \frac{\|\underline{A}_k\|_F}{\|\underline{A}\|_F} = \frac{\sqrt{\sigma_1^2+\sigma_2^2+\cdots+\sigma_k^2}}{\sqrt{\sigma_1^2+\cdots+\sigma_r^2}} \geq \alpha$$ 门限

选择满足 $\nu(k) \geq \alpha$ 的最小整数作为有效秩估计. α 是接近于1的阈值. $\alpha = 0.997$

图像、视频压缩.

$\underline{A}$: $n \times n$

$\underline{A}_k$ $\underline{U}_k \underline{\Sigma}_k \underline{V}_k$ $nk + k + kn = (2n+1)k$

nk k kn

压缩: $k < \frac{n^2}{2n+1}$

通常: $\underline{A}$ 矩阵稀疏

5.15: 判断有效秩的门限: (习题) 5.9. 5.10 5.14 5.15

$$\text{rank} = \max_i \frac{\sigma_i}{\sigma_1} = \varepsilon$$

$\varepsilon = 5\%$. 3%

第六章 (工程应用书) 矩阵方程求解

6.1 最小二乘方法:

1. 普通最小二乘方法:

超定方程：$\underset{m\times n}{\underline{\underline{A}}}\,\underline{x} = \underline{b}_{m\times 1}$ $\quad m>n \qquad \underline{b}+\Delta\underline{b} \doteq \underline{b}_0+\underline{e}+\Delta\underline{b} \to \underline{b}_0$

$\underline{b} = \underline{b}_0 + \underline{e}$ ↘ 加性观测误差或噪声

$\underline{\underline{A}}\,\underline{x} = \underline{b} + \Delta\underline{b}$ ↘ 校正向量

最小二乘的核心思想：求出的解向量 x 能够使矩阵方程两边的误差平方和最小

$\|\Delta\underline{b}\|_2^2 \to \min$

$\min\limits_{\underline{x}} \|\underline{\underline{A}}\,\underline{x} - \underline{b}\|_2^2$ 可以理解为 $\min\limits_{\underline{x}} \|\Delta\underline{b}\|_2^2$ 条件优化问题

s.t. $\underline{\underline{A}}\,\underline{x} = \underline{b} + \Delta\underline{b}$

↓

LS

不同应用领域：不同解释

目标函数：$\phi(\underline{x}) = (\underline{\underline{A}}\,\underline{x} - \underline{b})^H (\underline{\underline{A}}\,\underline{x} - \underline{b})$

$= \underline{x}^H \underline{\underline{A}}^H \underline{\underline{A}}\,\underline{x} - \underline{x}^H \underline{\underline{A}}^H \underline{b} - \underline{b}^H \underline{\underline{A}}\,\underline{x} + \underline{b}^H \underline{b}$

一阶导 $=0$. 二阶导 ≥ 0 这里可以不讲

$\underline{\underline{A}}^H \underline{\underline{A}}$ 必须正定 $\underline{\underline{A}}^H \underline{\underline{A}} \succeq 0$

信号处理：相关矩阵（或者估计）$\hat{\underline{\underline{R}}}_x \succ 0$

$\hat{\underline{\underline{R}}}_x + \delta \underline{\underline{I}} \succ 0$

正则化

δ 取值？一定对角矩阵

默认 $\underline{\underline{A}}^H \underline{\underline{A}}$ 正定.

$\dfrac{\partial \phi}{\partial \underline{x}} = (\underline{x}^H \underline{\underline{A}}^H \underline{\underline{A}})^T - (\underline{b}^H \underline{\underline{A}})^T = 0$

$\dfrac{\partial \phi}{\partial \underline{x}^*} = \underline{\underline{A}}^H \underline{\underline{A}}\,\underline{x} - \underline{\underline{A}}^H \underline{b} = 0$

$\underline{x} = (\underline{\underline{A}}^H \underline{\underline{A}})^{-1} \underline{\underline{A}}^H \underline{b}$ 秩缺：SVD. 伪逆 $\to \underline{x}_{LS} = (\underline{\underline{A}}^H \underline{\underline{A}})^{\dagger} \underline{\underline{A}}^H \underline{b}$

6.1.2 Gauss-Markov 定理：

特定条件下 LS 最优，使得 LS 应得了扩大

在参数估计理论中，称参数向量 θ 的估计 $\hat{\theta}$ 为无偏估计，若 $E(\hat{\theta})=\theta$（真实的未知参数向量）。如果一个无偏估计还具有最小方差，则称这一无偏估计为最优无偏估计

* 最优无偏解（统计信号处理）

须满足的条件：

$\underline{A}\,\underline{x}=\underline{b}+\underline{e}$ → 随机误差向量

$$\begin{cases}\text{误差均值为0} \quad E\{\underline{e}\}=0 \\ \text{协方差} \quad \operatorname{cov}\{\underline{e}\}=E\{\underline{e}\,\underline{e}^H\}=\sigma^2\underline{\underline{I}}\end{cases}$$

高斯白噪声

最优无偏解由最小二乘解给出

* 最大似然解 6.1.3

（概率）

3. LS 解与最大似然解的等价性（了解）

LS 被广泛应用

复高斯随机向量（加性误差向量 $\underline{e}=[e_1, e_2 \dots e_m]^T$ 为独立同分布）

$$f(\underline{e})=\frac{1}{\pi^m|\underline{\underline{\Gamma}}_e|}\exp[-(\underline{e}-\underline{u}_e)^H\underline{\underline{\Gamma}}_e^{-1}(\underline{e}-\underline{u}_e)]$$

概率密度函数

协方差矩阵：

误差分量的各分量均具有零均值和相同方差 σ^2，$\underline{\underline{\Gamma}}_e=\sigma^2\underline{\underline{I}}$

$$=\frac{1}{(\pi\sigma^2)^m}\exp[-\frac{1}{\sigma^2}\underline{e}^H\underline{e}]$$

$\|\underline{A}\underline{x}-\underline{b}\|^2$

当误差向量 $\underline{e}$ 为零均值的高斯随机向量，但其元素具有不同方差时，这种情况下的最大似然解将不可能等于最小二乘解公式，最小二乘解不再是最优的

估计 $\underline{x}$ 使得 $f(\underline{e})$ 最大 $\Leftrightarrow \min\|\underline{A}\underline{x}-\underline{b}\|_2^2$

4. 数据最小二乘

$\underline{A}=\underline{A}_0+\underline{E}$ $\qquad \underline{A}$：数据矩阵：观测误差或噪声

对 $\underline{A}$ 校准 $\quad (\underline{A}+\Delta\underline{A})\underline{x}=\underline{b}$