统计计算

2016 年 3 月 15 日

目录

目录

目录

目录

旋转变换

在平面 $\mathbb{R}^2$ 中将非零向量 $\mathbf{x}$ 顺时针旋转角度 θ 成为 $\mathbf{y}$ 的变换为:

$$\mathbf{y} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \mathbf{x}.$$

一般地, 设 $\mathbf{e}_1, \cdots, \mathbf{e}_n$ 为 Euclid 空间 $\mathbb{R}^n$ 的一组标准正交基, 可在平面 $(\mathbf{e}_i, \mathbf{e}_j)(i \neq j)$ 中定义旋转变换.

Givens 变换的定义

设实数 c 与 s 满足 $c^2+s^2=1$, 称

$$\mathbf{G}_{ij} = \begin{array}{cc}
\begin{bmatrix}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
& & & c & & & & s & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & -s & & & & c & & & \\
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{bmatrix} &
\begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix} \\
\begin{matrix} & & & i & & & & j & & & \end{matrix} &
\end{array}$$

为 Givens 矩阵 (初等旋转矩阵), 记为 $\mathbf{G}_{ij}(c,s)$.

Givens 变换的性质

1. Givens 矩阵为正交矩阵, 且
$$\mathbf{G}_{ij}(c,s)^{-1} = \mathbf{G}_{ij}(c,s)^{\mathrm{T}} = \mathbf{G}_{ij}(c,-s).$$
2. $\det \mathbf{G}_{ij}(c,s) = 1$.
3. 分块对角矩阵 $\mathrm{diag}\left(\mathbf{I}_m, \mathbf{G}_{ij}(c,s), \mathbf{I}_n\right)$ 为 Givens 矩阵.
4. 教材第 240 页性质 (2), (3).

引理

对任意非零向量 $\mathbf{x} \in \mathbb{R}^n (n > 1)$, 存在有限个 Givens 矩阵的乘积 $\mathbf{G}$, 使得 $\mathbf{Gx} = |\mathbf{x}|\mathbf{e}_1$, 其中 $\mathbf{e}_1 \in \mathbb{C}^n$ 为标准单位向量.

证明　记 $\mathbf{x} = [x_1, \cdots, x_n]^{\mathrm{T}}$. 先考虑 $x_1 \neq 0$ 的情形. 对 $\mathbf{x}$ 可构造 Givens 矩阵 $\mathbf{G}_{12}(c_1, s_1)$, 其中

$$c_1 = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad s_1 = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}.$$

对

$$\mathbf{G}_{12}\mathbf{x} = \left[\sqrt{x_1^2 + x_2^2}, 0, x_3, \cdots, x_n\right]^{\mathrm{T}}$$

构造 Givens 矩阵 $\mathbf{G}_{13}(c_2, s_2)$, 其中

$$c_2 = \frac{\sqrt{x_1^2 + x_2^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}}, \quad s_2 = \frac{x_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}}.$$

如此下去，可得 Givens 矩阵 $\mathbf{G}_{1,i+1}(c_i, s_i)$，$i = 1, \cdots, n-2$.

在 $n-2$ 步后对 $\mathbf{G}_{1,n-1} \cdots \mathbf{G}_{12}\mathbf{x}$ 构造 Givens 矩阵 $\mathbf{G}_{1n}(c_{n-1}, s_{n-1})$，其中

$$c_{n-1} = \frac{\sqrt{x_1^2 + \cdots + x_{n-1}^2}}{\sqrt{x_1^2 + \cdots + x_n^2}},$$

$$s_{n-1} = \frac{x_n}{\sqrt{x_1^2 + \cdots + x_n^2}}.$$

令

$$\mathbf{G} = \mathbf{G}_{1n}\mathbf{G}_{1,n-1}\cdots\mathbf{G}_{12},$$

则有

$$\mathbf{G}\mathbf{x} = \left[\sqrt{x_1^2+\cdots+x_n^2}, 0, \cdots, 0\right]^{\mathrm{T}} = |\mathbf{x}|\mathbf{e}_1.$$

若 $x_1 = \mathbf{0}$, 令 $k = \min\{i : x_i \neq 0\}$, 则 $1 < k \leq n$, 以上步骤从 $\mathbf{G}_{1k}$ 开始进行即可. $\square$

Givens 变换 (续)

定理

任意非奇异矩阵均可通过左连乘一系列 Givens 矩阵化为上三角矩阵.

具体证明留作习题.

Givens 变换将矩阵三角化算法

用 Givens 变换化矩阵为上三角矩阵的过程见教材第 241 页.

算法也分两种情形讨论:

1. $\mathbf{A}_{n\times m}$ 满列秩. 详见教材第 241–242 页算法 2.2.
2. $\mathrm{rank}\mathbf{A}_{n\times m} = r < \min(m, n)$ 时与 Householder 变换方法一样分析可得相应的算法.

Givens 变换的应用

在统计回归计算中有非常有趣的应用.

作业

1. 习题
   P. 292: 5.11, 5.12.
2. 详细分析教材中给出的正交 -三角分解算法.
3. 详细给出 Gram-Schmidt 算法及其修正中例子的计算步骤.
4. 用 Givens 变换计算在介绍 Householder 变换时所给的 QR 分解实例.

问　题?

目录

本节的任务

第 249 页本节引言部分，主要是获得更为简单的正交 -对角分解形式.

目录

特征值问题

设 $\mathbf{A} \in \mathbb{R}^{n\times n}$ 为对称矩阵. 特征值问题是求实数 λ 和非零向量 $\mathbf{x}$, 使得

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

λ 称为 $\mathbf{A}$ 的特征值, $\mathbf{x}$ 称为 $\mathbf{A}$ 的属于 λ 的特征向量.

对称阵的谱分解

定理

设 $\mathbf{A} \in \mathbb{R}^{n\times n}$ 为对称矩阵, 则 $\mathbf{A}$ 有如下的分解式:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}} = \mathbf{U}\,\mathrm{diag}(\lambda_1, \cdots, \lambda_n)\mathbf{U}^{\mathrm{T}}, \tag{1}$$

其中, $\lambda_1, \cdots, \lambda_n$ 是 $\mathbf{A}$ 的特征值 (也称为谱), $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_n]$ 为 n 阶正交阵, 其第 i 列 $\mathbf{u}_i$ 为 $\mathbf{A}$ 的属于特征值 λ_i 的特征向量.

式 (1) 所给 $\mathbf{A}$ 的分解式称为 $\mathbf{A}$ 的谱分解 (或特征值分解).

对称阵的谱分解 (续)

式 (1) 可以写成:

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^{\mathrm{T}}.$$

或

$$\mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{U} = \mathbf{D} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n). \tag{2}$$

式 (2) 可解释为 $\mathbf{A}$ 经正交变换后化为对角阵 $\mathbf{D}$. 以下考虑如何求正交矩阵 $\mathbf{U}$.

回忆前面的 Householder 变换或 Givens 变换将矩阵三角化的方法, 可以将矩阵 $\mathbf{A}$ 的正交 -三角分解 $\mathbf{A}=\mathbf{QR}$ 中的正交矩阵 $\mathbf{Q}$ 看成分解为一系列简单的 Householder 矩阵 (或 Givens 矩阵) 的乘积.

利用这一思想, 我们用迭代方法将分解式 (2) 中的 $\mathbf{U}$ 也分解为一系列简单或易于计算的正交矩阵的乘积.

对称阵谱分解的计算方法 (续)

考虑如下过程:

$$\begin{aligned} \mathbf{A}_0 &= \mathbf{A}, \\ \mathbf{A}_k &= \mathbf{U}_k^{\mathrm{T}} \mathbf{A}_{k-1} \mathbf{U}_k, \quad k = 1, 2, \cdots \end{aligned}$$

显然, 对任意 k 有

$$\mathbf{A}_k = (\mathbf{U}_1 \cdots \mathbf{U}_k)^{\mathrm{T}} \mathbf{A} (\mathbf{U}_1 \cdots \mathbf{U}_k). \tag{3}$$

对称阵谱分解的计算方法 (续)

问题

如何选取适当的 $\mathbf{U}_k$, 使得 $\mathbf{A}_k$ 在 k 充分大时接近于一个对角阵, 也就是使 (3) 接近于 (2)?

上述问题可以用 Jacobi 变换实现. 这里, $\mathbf{U}_k^{\mathrm{T}}(k=1,\cdots,n)$ 被选为 Givens 矩阵 $\mathbf{G}_{ij}(c,s)$, 于是式 (2) 成为:

$$\mathbf{A}_k=\mathbf{G}_{ij}\mathbf{A}_{k-1}\mathbf{G}_{ij}^{\mathrm{T}}. \tag{4}$$

若 $\mathbf{A}_k$ 比 $\mathbf{A}_{k-1}$ 更接近于对角阵, 则我们就能期望迭代公式产生的 $\mathbf{A}_k$ 趋于对角阵 (当 $k\to\infty$ 时). 基于这种想法选择 Givens 矩阵 $\mathbf{G}_{ij}(c,s)$, 这就需要确定 i,j,c,s.

设 $\mathbf{A}_{k-1} = (a_{ij})_{n\times n}$. 选择的 c 和 s 要使得经过 Jacobi 变换 (4) 后的 a_{ij} 和 a_{ji} 同时为 0.

可以推导得 (详见教材第 250 页):

$$c = \left(\frac{1}{2}\left(1 + \frac{y}{\sqrt{x^2+y^2}}\right)^2\right)^{1/2}, \quad s = \frac{x}{2c\sqrt{x^2+y^2}}, \tag{5}$$

其中, $x = 2a_{ij}$, $y = a_{ii} - a_{jj}$.

Jacobi 变换的性质

1. 保持对称性;
2. $a^*_{ij} = a^*_{ji} = 0$, 且第 i, j 行和 i, j 列 (标记为虚线) 之外数据不变;
3. 两条水平虚线上对应元素平方和不变 (除两条垂直虚线上数据外), 两条垂直虚线上类似;
4. 所有元素平方和不变;
5. 非对角元素中只是由 a_{ij} 和 a_{ji} 变换为 0, 因此, $E(\mathbf{A}^*) = E(\mathbf{A}) - 2a^2_{ij}$.

证明 详见教材第 251 页. □

Jacobi 算法

1. $\mathbf{A}_0 = \mathbf{A}$;
2. 对 $k = 1, 2, \cdots$, 计算 $\mathbf{A}_k = \mathbf{U}_k^{\mathrm{T}} \mathbf{A}_{k-1} \mathbf{U}_k$, 其中 $\mathbf{U}_k$ 是由式 (5) 确定的 Givens 矩阵 (此处称为 Jacobi 矩阵), 它使 $\mathbf{A}_{k-1}$ 的非对角线元素中绝对值最大者 (记为 a_{ij}) 变为 0.
   当 a_{ij} 为 0 (或接近于 0) 时, $\mathbf{A}_{k-1}$ 已经是对角阵 (或近似对角阵), 则迭代停止.

详细内容见第 252 页算法 3.1.

一般地, 通过有限次旋转变换是不能将 $\mathbf{A}$ 化成对角阵的, 但是可以证明如下定理:

定理

$\lim\limits_{k\to\infty} \mathbf{A}_k$ 是对角阵.

证明　见教材第 252 页定理 3.2. $\square$

Jacobi 算法 (续)

由上面的分析已经知道:

1. Jacobi 算法实际上是求实对称矩阵 $\mathbf{A}$ 的特征值和特征向量的算法;
2. $\lim_{k\to\infty} \mathbf{A}_k$ 的对角线元素就是 $\mathbf{A}$ 的特征值;
3. 令 $\mathbf{U} = \mathbf{U}_1\mathbf{U}_2\cdots\mathbf{U}_k\cdots = [\mathbf{u}_1, \cdots, \mathbf{u}_n]$, 则 $\mathbf{u}_1, \cdots, \mathbf{u}_n$ 为 $\mathbf{A}$ 的特征向量.

Jacobi 算法 (续)

经典 Jacobi 算法的优点和缺点. 见第 253 页中间.

Jacobi 算法的修正

1. 循环 Jacobi 算法: 第 253 页最后一段;
2. 过关 Jacobi 算法: 第 254 页.

问　题?

目录

考虑一般 (不必对称) 非奇异矩阵 $\mathbf{A}$.

在式

$$\mathbf{A}_k = \mathbf{U}_k^{\mathrm{T}} \mathbf{A}_{k-1} \mathbf{U}_k$$

的迭代计算过程中, 若取 $\mathbf{U}_k$ 为 $\mathbf{A}_{k-1}$ 的正交 -三角分解式 $\mathbf{A}_{k-1} = \mathbf{Q}_{k-1}\mathbf{R}_{k-1}$ 中的正交矩阵 $\mathbf{Q}_{k-1}$, 则可得有名的求矩阵特征值的 QR 算法.

非奇异矩阵的特征值和特征向量 (续)

求任意非奇异矩阵的特征值

1. $\mathbf{A}_0 = \mathbf{A}$;
2. 对 $k = 1, 2, \cdots$, 做 $\mathbf{A}_{k-1}$ 的 QR 分解

$$\mathbf{A}_{k-1} = \mathbf{Q}_{k-1}\mathbf{R}_{k-1},$$

令

$$\mathbf{A}_k = \mathbf{Q}_{k-1}^{\mathrm{T}}\mathbf{A}_{k-1}\mathbf{Q}_{k-1} = \mathbf{R}_{k-1}\mathbf{Q}_{k-1}.$$

上述算法的理由:

1. $\mathbf{A}_k$ 与 $\mathbf{A}$ 酉相似, 因此具有完全相同的特征值;
2. 可以证明 $\mathbf{A}_k$ "基本收敛" 于一个三角矩阵, 其对角线上的元素就是 $\mathbf{A}$ 的特征值.

进一步的问题

对于最一般矩阵, 是否存在前述的正交分解 (正交-对角分解)?

问　题?

目录

奇异值分解

奇异值分解是对最一般矩阵的正交 -对角分解形式.

奇异值分解是矩阵分析、现代数值分析最基本和最重要的工具之一.

设 $\mathbf{A} \in \mathbb{C}^{m\times n}$, 且 $r = \mathrm{rank}\mathbf{A}$. 因为 $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ 为 Hermite 半正定矩阵, 且

$$\mathrm{rank}(\mathbf{A}^{\mathrm{H}}\mathbf{A}) = \mathrm{rank}\mathbf{A} = r,$$

所以 $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ 的特征值可排列为

$$\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n.$$

称 $\sigma_i = \sqrt{\lambda_i}(i = 1, \cdots, n)$ 为 $\mathbf{A}$ 的奇异值, 其中 λ_i 为 $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ 的特征值.

非奇异矩阵的奇异值分解

定理

设 $\mathbf{A} \in \mathbb{C}^{n\times n}$ 非奇异, 则存在 n 阶酉阵 $\mathbf{U}_1$ 和 $\mathbf{U}_2$ 使得

$$\mathbf{A} = \mathbf{U}_1\mathbf{D}\mathbf{U}_2^{\mathrm{T}},$$

其中 $\mathbf{D} = \mathrm{diag}(d_1, \cdots, d_n)$ 且 $d_i > 0$ 为 $\mathbf{A}$ 的奇异值.

定理的证明详见教材第 255 页定理 3.3.

奇异值分解的存在性

设 $\mathbf{A} \in \mathbb{C}^{m \times n}$, 且 $r = \mathrm{rank}\mathbf{A}$, 则存在酉阵 $\mathbf{U} \in \mathbb{C}^{m \times m}$ 和 $\mathbf{V} \in \mathbb{C}^{n \times n}$, 使得

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} \Sigma & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \mathbf{V}^{\mathrm{H}},$$

其中 $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, 且 $\sigma_1 \geq \cdots \geq \sigma_r > 0$ 为 $\mathbf{A}$ 的奇异值. 上式称为 $\mathbf{A}$ 的奇异值分解.

证明　设 $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ 有谱分解

$$\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{V}\,\mathrm{diag}(\sigma_1^2, \cdots, \sigma_n^2)\mathbf{V}^{\mathrm{H}},$$

其中 $\mathbf{V}$ 为酉阵, 且

$$\sigma_1^2 \geq \cdots \geq \sigma_r^2 > 0 = \sigma_{r+1}^2 = \cdots = \sigma_n^2.$$

将酉阵 $\mathbf{V}$ 分块为 $\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2]$, 其中 $\mathbf{V}_1 \in \mathbb{C}^{n\times r}$, $\mathbf{V}_2 \in \mathbb{C}^{n\times(n-r)}$, 并令

$$\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_r),$$

于是

$$\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V}_1 = \mathbf{V}_1\Sigma^2, \tag{6}$$

$$\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V}_2 = \mathbf{O}. \tag{7}$$

令 $\mathbf{U}_1 = \mathbf{A}\mathbf{V}_1\Sigma^{-1}$, 则由式 (6) 知

$$\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1 = \Sigma^{-1}\mathbf{V}_1^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V}_1\Sigma^{-1} = \mathbf{I}.$$

可选择 $\mathbf{U}_2 \in \mathbb{C}^{m\times(m-r)}$, 使得 $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2]$ 为 m 阶酉阵.

再由式 (7) 及上述结论知

$$\begin{aligned} \mathbf{A}\mathbf{V}_2 &= \mathbf{O}, \\ \mathbf{A}\mathbf{V}_1 &= \mathbf{U}_1\Sigma, \\ \mathbf{U}_1^{\mathrm{H}}\mathbf{A}\mathbf{V}_1 &= \Sigma^{-1}\mathbf{V}_1^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{V}_1 = \Sigma. \end{aligned} \tag{8}$$

由式 (8) 可得

$$\mathbf{U}^{\mathrm{H}}\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{U}_1^{\mathrm{H}}\mathbf{A}\mathbf{V}_1 & \mathbf{U}_1^{\mathrm{H}}\mathbf{A}\mathbf{V}_2 \\ \mathbf{U}_2^{\mathrm{H}}\mathbf{A}\mathbf{V}_1 & \mathbf{U}_2^{\mathrm{H}}\mathbf{A}\mathbf{V}_2 \end{bmatrix}$$
$$= \begin{bmatrix} \Sigma & \mathbf{O} \\ \mathbf{U}_2^{\mathrm{H}}\mathbf{U}_1\Sigma & \mathbf{O} \end{bmatrix}$$
$$= \begin{bmatrix} \Sigma & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

定理证毕. $\square$

奇异值分解 (续)

上述定理的证明过程实际上给出了矩阵奇异值分解的一种算法.

值得注意的是: 尽管 $\mathbf{A}$ 的奇异值是唯一确定的, 但从定理的证明过程可以看出, 酉阵 $\mathbf{U}$ 和 $\mathbf{V}$ 一般是不唯一的, 因而矩阵 $\mathbf{A}$ 的奇异值分解一般不唯一.

矩阵 $\mathbf{A} \in \mathbb{C}^{m\times n}$ 奇异值分解的其它形式:

1. $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}}$, 其中 $\mathbf{U} \in \mathbb{C}^{m\times r}$ 和 $\mathbf{V} \in \mathbb{C}^{n\times r}$ 均为列正交矩阵, $\Sigma \in \mathbb{R}^{r\times r}$;
2. 令 $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_r]$, $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_r]$, 则

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

奇异值分解算法

算法 1: 利用定理证明过程

见教材第 257 页算法 3.2.

算法 2: 利用 Householder 变换

见教材第 257 页算法 3.3. 注意条件: $\mathbf{A}$ 满列秩.

算法 1: 利用定理证明过程

见教材第 257 页算法 3.2.

算法 2: 利用 Householder 变换

见教材第 257 页算法 3.3. 注意条件: $\mathbf{A}$ 满列秩.

作业

习题　P. 293: 5.13, 5.14.

奇异值分解的应用

奇异值分解在数据压缩中具有很好的作用.

设 $\mathbf{A} \in \mathbb{C}^{m\times n}$, 且 $r = \operatorname{rank}\mathbf{A} > 1$, 我们希望寻求一个低秩矩阵 $\mathbf{B}$ 以逼近原矩阵 $\mathbf{A}$, 即考虑如下优化问题

$$\min_{\operatorname{rank}\mathbf{B}=k} \|\mathbf{A}-\mathbf{B}\|, \quad k < r.$$

当然, 上述优化问题的解依赖于矩阵范数的选择, 常用谱范数和 Frobenius 范数.

利用矩阵的奇异值分解可以非常方便地求出上述逼近问题的最优低秩矩阵.

设 $\mathbf{A}$ 有奇异值分解 $\mathbf{A} = \mathbf{U}\,\mathrm{diag}(\Sigma, \mathbf{O})\mathbf{V}^{\mathrm{H}}$, 其中 $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_m]$ 和 $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_n]$ 均为酉阵, $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, 且 $\sigma_1 \geq \cdots \geq \sigma_r > 0$ 为 $\mathbf{A}$ 的奇异值, 即

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

对任意正整数 $k < r$, 令

$$\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{H}}.$$

奇异值分解的应用 (续)

显然 $\mathrm{rank}\mathbf{A}_k = k$, 且可证明

$$\min_{\mathrm{rank}\mathbf{B}=k} \|\mathbf{A}-\mathbf{B}\|_2 = \|\mathbf{A}-\mathbf{A}_k\|_2 = \sigma_{k+1},$$

$$\min_{\mathrm{rank}\mathbf{B}=k} \|\mathbf{A}-\mathbf{B}\|_F = \|\mathbf{A}-\mathbf{A}_k\|_F = \left(\sum_{i=k+1}^{r} \sigma_i^2\right)^{1/2}.$$

问　题?