回顾 对称矩阵正交对角化.

$A = A^T \in M_n(\mathbb{R})$, $\exists\, Q \in O(n)$. s.t. $Q^T A Q = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$

$\lambda_i \in \mathbb{R}$, $(Q^{-1} A Q)$

(相合, 相似)

也称作自伴线性变换的谱定理. $T: V \to V$. V是内积空间.
T自伴, 则T在标准正交基下可对角化 (V分解成一维T不变子空间的正交直和)

结合正定性判定: $A = A^T$ 正定 $\Leftrightarrow \lambda_i > 0$
$\Leftrightarrow A = P \cdot P^T$, P可逆.

$A = A^T$ 半正定 $\Leftrightarrow \lambda_i \geq 0 \Leftrightarrow A = P \cdot P^T$.

如果 $T: V \to W$ 是两个内积空间之间的线性映射.
如何找到T的标准型?

$A \in M_{m\times n}(\mathbb{R})$, 回顾 $\exists\, P \in GL(n, \mathbb{R})$
$Q \in GL(m, \mathbb{R})$

s.t. $Q^{-1} A P = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$

$r = \mathrm{rk}\, A$. (相抵标准型)

现在附加内积结构, 要求 P. Q 正交.

Thm (SVD 奇异值分解)

$A \in M_{m\times n}(\mathbb{R})$. $\exists\ P \in O(n)$. $Q \in O(m)$,

s.t. $A = Q D P^T$ $(P^T = P^{-1})$

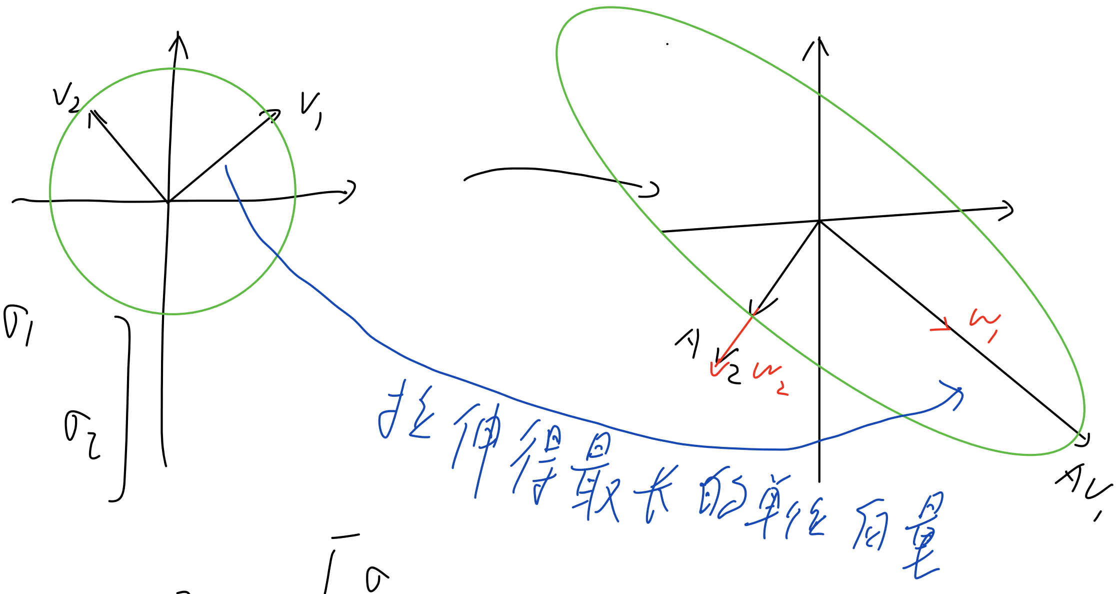
$$D = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ \hline & 0 & \end{bmatrix}$$

$\sigma_1 \geq \sigma_2 \cdots \geq \sigma_n \geq 0$，由 A 唯一确定，称作 A 的奇异值.

几何解释:

$A(v_1, v_2)$

$= (w_1, w_2)\begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix}$



($m \leq n$ 时. $D = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_n & \end{array}\right]$)

(只需考虑 $m \geq n$，另一半对 A 取转置即可)

若 $\sigma_1 \geq \cdots \geq \sigma_r > 0$. $\sigma_{r+1} = 0$, $r = \mathrm{rk}\, A$.

$Q = (w_1 \cdots w_m)$, $P = (v_1 \cdots v_n)$. $w_1 \cdots w_m$ $\mathbb{R}^m$ 标准正交基

$A = \sum_{i=1}^{r} \sigma_i w_i \cdot v_i^T$ (常用形式) $v_1 \cdots v_n$ $\mathbb{R}^n$ 标准正交基

证明：考虑 A^TA，对称，半正定。

Step 1：$\sigma_1 \cdots \sigma_n$ 唯一性

若 $A = QDP^T$，则 $A^TA = P \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix} P^T$，

所以 $\sigma_1^2 \cdots \sigma_n^2$ 是 A^TA 的特征值，

排序后由 A 唯一确定。

Step 2：SVD 存在性（找 P，Q）

想法：
由 Step 1，P 应该出现在 A^TA 的正交对角化中。

A^TA 对称 $\Rightarrow \exists P \in O(n)$, s.t.

$$P^{-1}(A^TA)P = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad \lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_n \geq 0$$

$P = (v_1 \cdots v_n)$.

$A^TAv_i = \lambda_i v_i$.

Claim: $Av_1 \cdots Av_n$ 相互正交.

$i \neq j$, $\langle Av_i, Av_j \rangle_{\mathbb{R}^m} = (Av_i)^T (Av_j)$

$= v_i^T (A^T A v_j)$

$= \lambda_j v_i^T v_j$

$= \lambda_j \langle v_i, v_j \rangle_{\mathbb{R}^n} = 0$

$i = j$, $\Rightarrow \langle Av_i, Av_i \rangle_{\mathbb{R}^m} = \lambda_i \langle v_i, v_i \rangle_{\mathbb{R}^n} = \lambda_i$

假设 $\lambda_1 \geq \cdots \geq \lambda_r > 0$, $\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$

取 $w_1 = \dfrac{Av_1}{|Av_1|_{\mathbb{R}^m}} \cdots w_r = \dfrac{Av_r}{|Av_r|_{\mathbb{R}^m}}$,

$w_1, \cdots, w_r$ 标准正交, 扩充为标准正交基,

则 $|A \cdot v_i|_{\mathbb{R}^m} = \sqrt{\lambda_i} \overset{\text{记为}}{=} \sigma_i$

$$(Av_1 \cdots Av_n) = (w_1 \cdots w_m) \cdot \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{bmatrix}}_{D}$$

即 $A \cdot P = Q \cdot D$, $P \in O(n)$, $Q \in O(m)$

应用场景 (1) 图像压缩.

$A=(a_{ij})_{m\times n}$ a_{ij} 表达 i,j 点 像素的灰度.

m,n 很大, 记忆 mn 数据.

$$A=\sum_{i=1}^{\min(m,n)}\sigma_i w_i v_i^T, \quad \text{取 } k \ll \min(m,n)$$

$$A_k=\sum_{i=1}^{k}\sigma_i w_i v_i^T, \quad \text{记忆 } k(m+n+1) \text{ 数据.}$$

$$\left(\begin{array}{ccc|c} w_1 & \cdots & w_k & 0 \end{array}\right)\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}\begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \\ \hline 0 \end{bmatrix}$$

A_k 很接近 A.

(什么意义下)

(Brunton-Kutz 书中的图像)

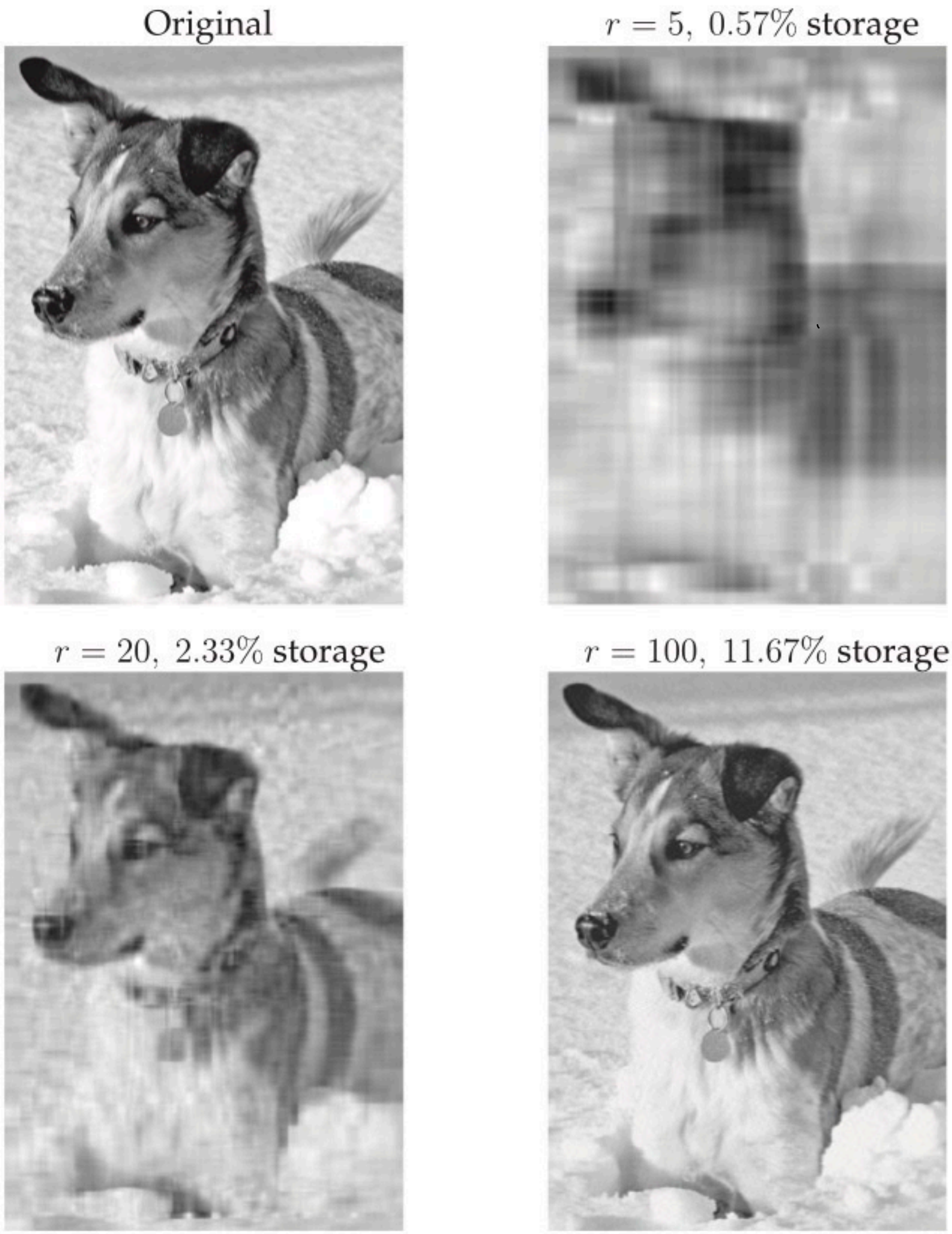


Figure 1.3 Image compression of Mordecai the snow dog, truncating the SVD at various ranks r. Original image resolution is 2000×1500.

Brunton-Kutz, Data-driven Science and engineering
Page 10

$M_{m\times n}(\mathbb{R})$上有内积 $\langle\ ,\ \rangle$.（Frobenius 内积）

$\langle A, B\rangle = Tr(A^T B)$, $\langle A, A\rangle = \|A\|_F$

则 A, B 距离 $\sqrt{\langle A-B, A-B\rangle}$

$= \sqrt{\sum_{i,j}(a_{ij}-b_{ij})^2}$ （与直观相符

等同于 $\mathbb{R}^{mn}$ 的标准

内积.）

则定理（Eckart-Young, Schmidt）（低秩逼近）

A_k 是 $rk \le k$ 的 $m\times n$ 矩阵中与 A 距离最近

的

$$\|A-A_k\|_F = \min_{rkB\le k}\|A-B\|$$

差多少呢??

引理(HW). $V = M_{m\times n}(\mathbb{R})$, $\langle , \rangle_F$

则 $\forall P \in O(n)$, $Q \in O(m)$.

$T_{P,Q}: V \longrightarrow V$ 是正交变换.

$A \longmapsto QAP$

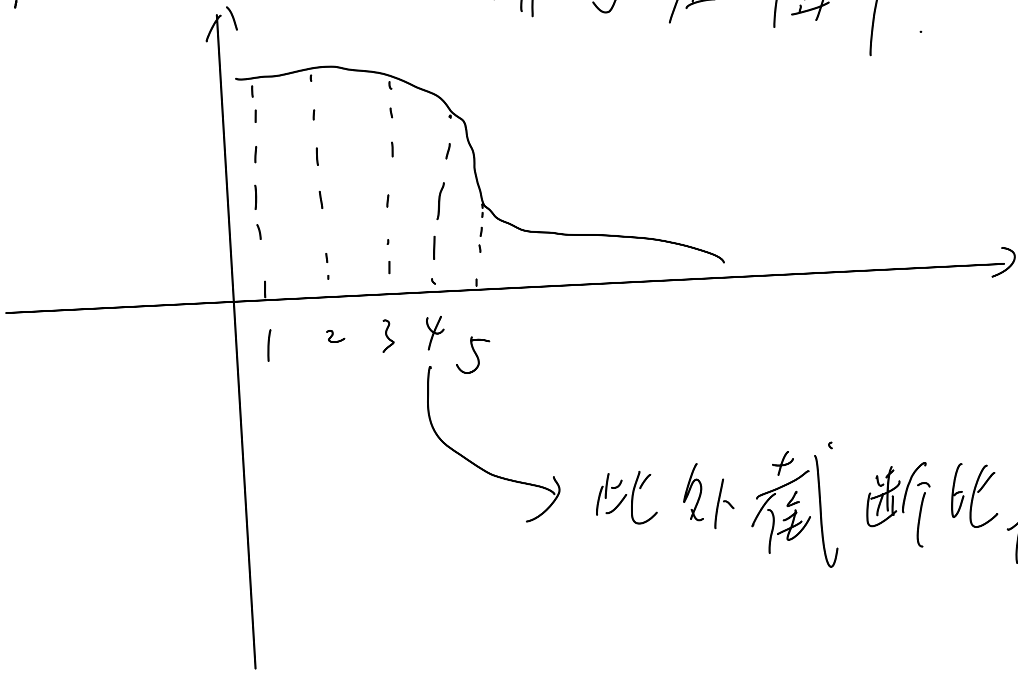
推论1: $A = QDP^T$, $D = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & \end{bmatrix}$

则 $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_n^2}$

推论2:

$A_k = Q \cdot \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} P^T$, $\|A - A_k\|_F = \underbrace{\sqrt{\sigma_{k+1}^2 + \cdots + \sigma_r^2}}_{\text{也称为 lost energy}}$

将 $r_1, \ldots, r_k$ 排列在图中.



此外截断比较经济.