RWTH AACHEN UNIVERSITY

2. Prerequisites from Matrix Algebra

Real $(m \times n)$ matrices will be written as

$$M = (m_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathbb{R}^{m \times n} \quad (\text{or } \mathbb{C}^{m \times n})$$

Diagonal matrices as $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n \times n}$

Matrix $U \in \mathbb{R}^{n \times n}$ is called orthogonal if

$$UU^T = U^T U = I_n \quad (I_n \text{ identity matrix})$$

$\mathcal{O}(n)$ denotes the set of orthogonal matrices.

Th. 2.1 (Singular Value Decomposition, SVD)

Given $M \in \mathbb{R}^{m \times n}$. There exist $U \in \mathcal{O}(m)$ and $V \in \mathcal{O}(n)$ and some $\Sigma \in \mathbb{R}^{m \times n}$ with non-negative entries in its diagonal and zeros otherwise such that

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

The diagonal values of Σ are called singular values.

The columns of U and V are called left and right singular ~~values~~ vectors.

Thinking the Future
Zukunft denken

Remark: If $m<n$, say, SVD may be written as

$\exists\, U \in \mathbb{R}^{m\times n}$, $UU^T = I_m$, $\exists\, V \in \mathcal{O}(n)$, $\exists\, \Sigma \in \mathbb{R}^{n\times n}$ diagonal:

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

Th. 2.2. (Spectral decomposition)

Given $M \in \mathbb{R}^{n\times n}$ symmetric. There $\exists\, V \in \mathcal{O}(n)$, $V = (v_1, \dots, v_n)$ and $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ such that

$$M = V\Lambda V^T = \sum_{i=1}^{n} \lambda_i v_i v_i^T$$

v_i are the eigenvectors of M with eigenvalues λ_i. ⌟

- If $\lambda_i > 0$, $i = 1, \dots, n$, M is called positive definite p.d. $(M > 0)$

  If $\lambda \geq 0$, $i = 1, \dots, n$, M is called non-negative definite n.n.d. $(M \geq 0)$

- If M is n.n.d., then it has Cholesky decomposition

$$M = V\Lambda^{1/2} (V\Lambda^{1/2})^T$$

  where $\Lambda^{1/2} = \operatorname{diag}(\lambda_1^{1/2}, \dots, \lambda_n^{1/2})$.

- $M \geq 0 \iff x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n$

- $M > 0 \iff x^T M x > 0 \quad \forall x \in \mathbb{R}^n, x \neq 0$.

D. 2.3. a) Given $M \in \mathbb{R}^{4\times 4}$, $M = (m_{ij})_{1\leq i,j\leq 4}$.

$tr(M) = \sum_{i=1}^{4} m_{ii}$ called trace of M.

b) Given $M \in \mathbb{R}^{4\times 4}$. $\|M\|_F = \sqrt{\sum_{i,j} m_{ij}^2} = \sqrt{tr(M^T M)}$

is called the Frobenius norm of M.

c) $M \in \mathbb{R}^{4\times 4}_{sym}$, M symmetric. $\|M\|_s = \max_{1\leq i\leq 4} |\lambda_i|$ is called spectral norm.

(A "norm" for symmetric matrices.)

- It holds that $tr(A \cdot B) = tr(B \cdot A)$, $A \in \mathbb{R}^{m\times 4}$, $B \in \mathbb{R}^{4\times m}$.
- $tr(M) = \sum_{i=1}^{4} \lambda_i(M)$, $det(M) = \prod_{i=1}^{4} \lambda_i$, M symm.

$tr(M) = tr(V \Lambda V^T) = tr(\Lambda \underbrace{V^T V}_{I_4}) = tr(\Lambda) = \sum_{i=1}^{4} \lambda_i(M)$

$det(M) = det(V \Lambda V^T) = \underbrace{det(V)}_{=1} \cdot det \Lambda \cdot \underbrace{det V^T}_{=1} = det \Lambda$

$= \prod_{i=1}^{4} \lambda_i(M)$.

<u>Th. 2.4.</u> (Ky Fan, 1950)

Given $M \in \mathbb{R}^{n\times n}$ symm., $k \le n$, $\lambda_1(M) \geq \dots \geq \lambda_n(M)$ eigenvalues.

$$\max_{\substack{V\in\mathbb{R}^{n\times k}\\ V^TV=I_k}} \operatorname{tr}(V^TMV) = \sum_{i=1}^{k} \lambda_i(M)$$

$$\min_{\substack{V\in\mathbb{R}^{n\times k}\\ V^TV=I_k}} \operatorname{tr}(V^TMV) = \sum_{i=1}^{k} \lambda_{n-i+1}(M)$$

$\perp$

∘ Special case of Th. 2.4. $(k=1)$

$$\max_{\|v\|=1} v^TMv = \lambda_{max}(M)$$

$$\min_{\|v\|=1} v^TMv = \lambda_{min}(M)$$

Note that $\displaystyle \max_{\|v\|=1} v^TMv = \max_{v\neq 0} \frac{v^TMv}{v^Tv}$

<u>Th. 2.5</u> Given $A, B \in \mathbb{R}^{n\times n}_{sym}$ with eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$, $\mu_1 \geq \dots \geq \mu_n$, respectively. Then

$$\sum_{i=1}^{n} \lambda_i \mu_{n-i+1} \le \operatorname{tr}(A\cdot B) \le \sum_{i=1}^{n} \lambda_i \mu_i$$

$\perp$

Let $\lambda^+ = \max\{\lambda, 0\}$ positive part of $\lambda \in \mathbb{R}$.

<u>Th. 2.6.</u> Given $M \in \mathbb{R}^{n\times n}$ symmetric with spectral decomp.
$M = V \operatorname{diag}(\lambda_1, \dots, \lambda_n) V^T$, $\lambda_1 \geq \dots \geq \lambda_n$. Then

$$\min_{A \succeq 0,\ \operatorname{rk}(A) \leq k} \|M - A\|_F^2$$

is attained at $A^* = V \operatorname{diag}(\lambda_1^+, \dots, \lambda_k^+, 0, \dots, 0) V^T$

with optimum value $\sum_{i=1}^{k} (\lambda_i - \lambda_i^+)^2 + \sum_{i=k+1}^{n} \lambda_i^2$. $\lrcorner$

<u>Proof.</u>

$$\|M - A\|^2 = \operatorname{tr}\left[(M-A)(M-A)^T\right]$$
$$= \|M\|^2 - 2\operatorname{tr}(MA) + \|A\|^2$$
$$\geq \sum_{i=1}^{n} \lambda_i^2 - 2 \sum_{i=1}^{n} \lambda_i \mu_i + \sum_{i=1}^{n} \mu_i^2$$
$$= \sum_{i=1}^{n} (\lambda_i - \mu_i)^2$$
$$= \sum_{i=1}^{k} (\lambda_i - \mu_i)^2 + \sum_{i=k+1}^{n} \lambda_i^2$$
$$\geq \sum_{i=1}^{k} (\lambda_i - \lambda_i^+)^2 + \sum_{i=k+1}^{n} \lambda_i^2$$

$\mu_1 \geq \dots \geq \mu_n \geq 0$ eigenvalues of A

$\mu_1, \dots, \mu_k$ may be positive

$\mu_{k+1}, \dots, \mu_n = 0$.

Lower bound is achieved if $A = V \operatorname{diag}(\lambda_1^+, \dots, \lambda_k^+, 0, \dots, 0) V^T$ $\blacksquare$

RWTH AACHEN UNIVERSITY

D.2.7 (Löwner semi-ordering)

Given $V, W \geq 0$. Define $V \leq W$ if $W - V \geq 0$ (n.n.d.)

Show that "$\leq$" fulfills

(i) $V \leq V$ ✓ (ii) $V \leq W$ and $W \leq V \Rightarrow V = W$

(iii) $U \leq V$ and $V \leq W \Rightarrow U \leq W$ ✓

Th. 2.8. Given V, W n.n.d., $V = (v_{ij})$, $W = (w_{ij})$

$\lambda_1(V) \geq \dots \geq \lambda_n(V)$, $\lambda_1(W) \geq \dots \geq \lambda_n(W)$.

~~a)~~ $\cancel{\Leftrightarrow}$ $V \leq W$

a) $\lambda_i(V) \leq \lambda_i(W)$, $i = 1, \dots, n$

b) $v_{ii} \leq w_{ii}$, $i = 1, \dots, n$

c) $v_{ii} + v_{jj} - 2v_{ij} \leq w_{ii} + w_{jj} - 2w_{ij}$.

d) $tr(V) \leq tr(W)$

e) $det(V) \leq det(W)$

b) $e_i^T V e_i \leq e_i^T W e_i$

$v_{ii} \leq w_{ii}$ ✓

Proof (Ex.)

Thinking the Future
Zukunft denken

<u>Def. 2.9.</u> $Q \in \mathbb{R}^{n \times n}$ is called <u>projection</u> (<u>matrix</u>), (or idempotent) if $Q^2 = Q$.
It is called <u>orthogonal projection</u> if additionally $Q = Q^T$.

Concept: Q maps onto $\mathrm{Im}(Q)$, a k-dim. subspace
Let $x \in \mathbb{R}^n$, $y = Qx \in \mathrm{Im}(Q)$; $Qy = y$
(no change anymore)

Orthogonality: $x \in \mathbb{R}^n$, $y = Qz \in \mathrm{Im}(Q)$

$$y^T (x - Qx) = z^T Q^T (x - Qx)$$
$$= z^T (Qx - Q^2 x) = z^T 0 = 0$$