$M \in \mathbb{R}^{n \times n}$ symm.,

Spectral decomposition

$$M = V \operatorname{diag}(\lambda_1, \dots, \lambda_n) V^T, \quad \lambda_1 \geq \dots \geq \lambda_n$$

$$V = (v_1, \dots, v_n), \quad V^T V = I_n$$

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

max is attained at $V^* = (v_1, \dots, v_k)$

(2) $$\min_{A \succeq 0,\ \operatorname{rk}(A) \leq k} \| M - A \|_F^2 = \sum_{i=1}^{k} (\lambda_i - \lambda_i^+)^2 + \sum_{i=k+1}^{n} \lambda_i^2$$

~~max~~ min attained at $\cancel{A^* = V \operatorname{diag}(\lambda_1^+, \dots, \lambda_k^+, \lambda_{k+1}, \dots, \lambda_n) V^T}$

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

Thinking the Future
Zukunft denken

<u>Def. 2.7.</u> (Löwner semi-ordering)

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

Show '$\leq$' is a semi-ordering on the set of n.n.d. mat.

(i) $V \leq V$ (reflexiv)

(ii) $V \leq W$ and $W \leq V \Rightarrow V = W$ (antisymmetric)

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

<u>Th. 2.8.</u> Given V, W n.n.d., $V = (v_{ij}) \leq W = (w_{ij})$, $\lambda_1(V) \geq \dots \geq \lambda_n(V)$, $\lambda_1(W) \geq \dots \geq \lambda_n(W)$ eigenvalues.

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

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

c) $v_{ii} + v_{jj} - 2v_{ij} \leq w_{ii} + w_{jj} - 2w_{ij}$ $\forall i,j = 1, \dots, n$

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

e) $det(V) \leq det(W)$ ⌟

(Proof as exercise)

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

It is called <u>orthogonal projection</u> if additionally $Q = Q^T$, i.e., Q is symmetric. ⌋

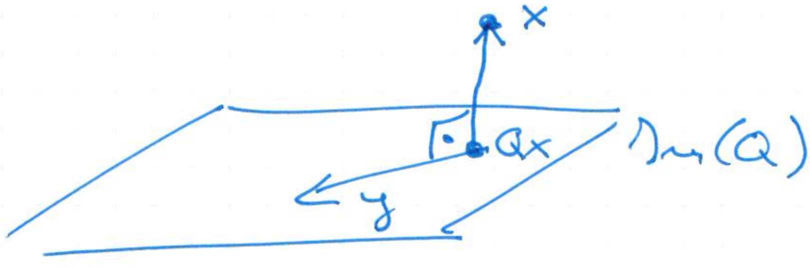
Concepts: Q maps onto $\operatorname{Im} Q$, a k-dim subspace.

Let $x \in \mathbb{R}^n$, $y = Qx \in \operatorname{Im}(Q)$, $Qy = Q^2 x = y$

(no change anymore)

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

$$
\begin{aligned}
y^T(x - Qx) &= z^T Q^T (x - Qx) \\
&= z^T Q (x - Qx) = z^T (Qx - Q^2 x) \\
&= z^T (Qx - Qx) = 0
\end{aligned}
$$



Th. 2.10 $V \in \mathbb{R}^{n \times k}$, $V = [v_1, \dots, v_k]$, $V^T V = I_k$, $k \le n$.

Then $Q = \sum_{i=1}^{k} v_i v_i^T$ is an orthogonal proj. onto $\operatorname{Im}(Q) = \langle v_1, \dots, v_k \rangle$. ⊥

Proof. $x \in \mathbb{R}^n$. $Qx = \sum_{i=1}^{k} v_i v_i^T x$

$$= \sum_{i=1}^{k} (v_i^T x) v_i = \sum_{i=1}^{k} y_i v_i \in \operatorname{Im}(Q)$$

$$Q^2 = \left(\sum_{i=1}^{k} v_i v_i^T\right)\left(\sum_{i=1}^{k} v_i v_i^T\right) = \sum_{i=1}^{k} v_i v_i^T = Q$$

and Q is symmetric. ∎

- Let Q be a orth. projection on $\operatorname{Im}(Q)$.

  Then $I-Q$ is an orth. proj. onto $\mathcal{K}(Q)$.

  [$\operatorname{Im}(Q)$ denotes the image of Q.
  $\mathcal{K}(Q)$ denotes the kernel / nullspace of Q.]

  Since: $(I-Q)^2 = (I-Q)(I-Q) = I - Q - Q + Q^2$

  $$= I - Q - Q + Q = I - Q.$$

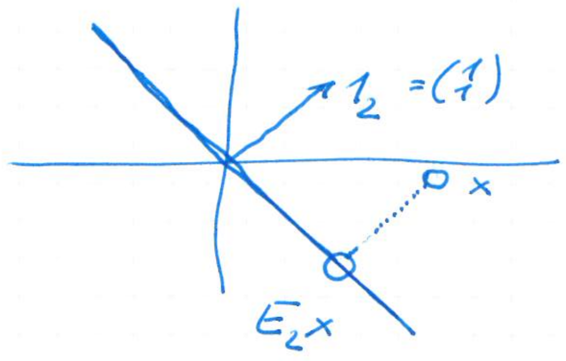
  Let $y \in \mathcal{K}(Q)$, i.e., $Qy = 0$. Then

  $$(I-Q)y = y - Qy = y \in \operatorname{Im}(I-Q).$$

- Define $E_n = I_n - \frac{1}{n} 1_{n\times n} = \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{pmatrix}$

  E_n is an orth. proj. onto $1_n^\perp = \{x \in \mathbb{R}^n / 1_n^T x = 0\}$

Since $1_4^T E_4 x = 1_4^T (I_4 - \frac{1}{4} 1_{4\times 4}) x$
$= (1_4^T - 1_4^T) x = 0 \quad \forall x \in \mathbb{R}^4$



Let $x \in \mathbb{R}^4$. $E_4 x = x - \bar{x}$, $\bar{x} = \frac{1}{4}\sum_{i=1}^{4} x_i$

$E_4 x$ centering at 0

$y = E_4 x$ satisfies $\sum_{i=1}^{4} y_i = 0$.

Th. 2.11. (Inverse & det of partitioned matrices)

Let $M = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ be symmetric, regular, A regular.

a) $M^{-1} = \begin{pmatrix} A^{-1} + FE^{-1}F^T & -FE^{-1} \\ -E^{-1}F^T & E^{-1} \end{pmatrix}$,

$E = C - B^T A^{-1} B$
(Schur complement)

$F = A^{-1} B$

b) $\det M = \det A \cdot \det(C - B^T A^{-1} B)$. $\perp$

Extension to general $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ exists, s., Murphy p. 118.

Isometry:

Def. 2.12. Matrix $M \in \mathbb{R}^{4\times4}$ is called isometry if $(Mx)^T(Mx) = x^Tx \quad \forall x \in \mathbb{R}^4$. $\perp$

Equivalent: $\|Mx\| = \|x\|$

Properties:

a) U, V isometries $\Rightarrow$ UV is isometry

b) U isometry $\Rightarrow$ $\det(U) = 1$

c) U isometry $\Rightarrow$ $|\lambda(U)| = 1$ for all eigenvalues of U.

Because:

a) $\|UVx\| = \|Vx\| = \|x\|$

c) λ eigenvalue of U, x corr. eigenvector

$Ux = \lambda x \Rightarrow \|Ux\| = \|x\| = |\lambda|\,\|x\| \Rightarrow |\lambda| = 1$

b) $\det U = \prod_{i=1}^{4} \lambda_i(U) = 1$ $\boxtimes$

3. Multivariate Distributions and Moments

3.1. Random vectors

Let $X_1, \dots, X_p$ be random variables on the same probability $(\Omega, \mathcal{A}, P)$: $X_i : (\Omega, \mathcal{A}, P) \to (\mathbb{R}^1, \mathcal{B}^1)$

- $X = (X_1, \dots, X_p)^T$ is called a random vector.
- Analogously: $X = (X_{ij})_{\substack{1 \le i \le p \\ 1 \le j \le n}}$ composed of r.v. X_{ij} is called a random matrix.
- The (joint) distribution of a random vector is uniquely described by its multivariate distribution function.

$$F(x_1, \dots, x_p) = P(X_1 \le x_1, \dots, X_p \le x_p), \quad (x_1, \dots, x_p) \in \mathbb{R}^p$$

- A random vector $X = (X_1, \dots, X_p)^T$ is called absolutely-continuous if there exists an integrable function $f(x_1, \dots, x_p) \ge 0$ s.t.

$$F(x_1, \dots, x_p) = \int_{-\infty}^{x_p} \dots \int_{-\infty}^{x_1} f(u_1, \dots, u_p) \, du_1 \dots du_p$$

f is called (probability) density function (pdf)
F is called cumulative distribution fct. (cdf).

Example 3.1.~~4~~ (Multivariate normal distr. / Gaussian)

It has a pdf

$$f(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right\},$$

$$x = (x_1, \dots, x_p) \in \mathbb{R}^p$$

with parameters $\mu \in \mathbb{R}^p$, $\Sigma \in \mathbb{R}^{p\times p}$, $\Sigma > 0$.

Abbr. $X = (X_1, \dots, X_p)^T \sim N_p(\mu, \Sigma)$

($\sim$: distributed as / according to ...)

Note: Σ must have full rank.
There exist a multivariate Gaussian, even if ~~Σ~~ $rk(\Sigma) < p$. But there is no pdf w.r.t. λ^p.