

Friederike Horn & Bileam Scheuven

Justify all your claims.

Exercise 1 (Principal Component Analysis)

- a) Suppose λ is an eigenvalue to the eigenvector v of Σ .

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle$$

$$= \langle \Sigma x, x \rangle$$

Since $\Sigma = \bar{\Sigma}$

$$= \langle x, \Sigma x \rangle$$

$$= \langle x, \lambda x \rangle$$

$$= \bar{\lambda} \langle x, x \rangle$$

If $\lambda = \bar{\lambda}$ then λ must be in \mathbb{R} .

- b) By the Rayleigh coefficient, vector that maximizes the variance is the normed eigenvector $\frac{v_1}{\|v_1\|}$ corresponding to the largest eigenvalue λ_1 of $X^T X$. The resulting variance is $\sqrt{\lambda_1}$ i.e. the first singular value σ_1 of X .
- c) Analogously, the vector orthogonal to v_1 that maximizes the variance is the normed eigenvector $\frac{v_2}{\|v_2\|}$ corresponding to the largest eigenvalue λ_2 of $X^T X$. The resulting variance is the second singular value σ_2 of X .
- d) The k -th principal component is obtained by eigendecomposition of $X^T X$ and ordering the eigenvalues such that $\lambda_1 < \lambda_2 < \dots < \lambda_d$. Then the k -th principal component is $X * v_k$ with variance $\sqrt{\lambda_k}$. To obtain the lower dimensional representation \tilde{X} we multiply the matrix with the eigenvectors of $X X^T$ as columns \tilde{U} by the matrix with diagonal elements $\tilde{\sigma}_i = \sqrt{\lambda_i}$ $\tilde{\Sigma}$ and the matrix with eigenvectors of $X^T X$ as rows \tilde{V}^T .

$$\tilde{X} = \tilde{U} \tilde{\Sigma} \tilde{V}^T$$

Exercise 2 (Singular Value Decomposition)

- a) The singular values are equal to the ordered eigenvalues of MM^T and $M^T M$.

$$M^T M = \begin{bmatrix} -4 & 1 & 3 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} -4 & 2 \\ 1 & 2 \\ 3 & 2 \end{bmatrix} = \begin{pmatrix} 26 & 0 \\ 0 & 12 \end{pmatrix}$$

Since the result is diagonal, the eigenvalues are easily determined as $\lambda_1 = 26, \lambda_2 = 12$ which results in the singular values $\sigma_1 = \sqrt{26}, \sigma_2 = \sqrt{12}, \sigma_3 = 0$.

- b) The singular values of a matrix A are the union of the roots of the eigenvalues of $A^T A$ and AA^T . Since the union is commutative, $eigs(AA^T) \cup eigs(A^T A) = eigs(A^T A) \cup eigs(AA^T)$, so the singular values are identical.

c) For $\Sigma^\# \Sigma \Sigma^\# = \Sigma^\#$ since Σ is diagonal, analogously for Σ .

•

$$\begin{aligned} AA^\# A &\stackrel{!}{=} A \\ U\Sigma V^T V\Sigma^\# U^T U\Sigma V^T &= A \\ U\Sigma \Sigma^\# \Sigma V^T &= A \\ U\Sigma V^T &= A \end{aligned}$$

•

$$\begin{aligned} A^\# AA^\# &\stackrel{!}{=} A^\# \\ V\Sigma^\# U^T U\Sigma V^T V\Sigma^\# U^T &= A^\# \\ V\Sigma^\# \Sigma \Sigma^\# U^T &= A^\# \\ V\Sigma^\# U^T &= A^\# \end{aligned}$$

•

$$\begin{aligned} (AA^\#)^T &\stackrel{!}{=} AA^\# \\ (U\Sigma V^T V\Sigma^\# U^T)^T &= AA^\# \\ U\Sigma^\# V^T V\Sigma U^T &= AA^\# \\ U\Sigma^\# V^T V\Sigma U^T &= U\Sigma V^T V\Sigma^\# U^T \\ U\Sigma^\# \Sigma U^T &= U\Sigma \Sigma^\# U^T \end{aligned}$$

•

$$\begin{aligned} (A^\# A)^T &\stackrel{!}{=} A^\# A \\ (V\Sigma^\# U^T U\Sigma V^T)^T &= A^\# A \\ V\Sigma U^T U\Sigma^\# V^T &= A^\# A \\ V\Sigma U^T U\Sigma^\# V^T &= V\Sigma^\# U^T U\Sigma V^T \\ V\Sigma \Sigma^\# V^T &= V\Sigma^\# \Sigma V^T \end{aligned}$$

Exercise 3 (Condition Number)

a) First prove that $\|Av\| \stackrel{!}{\leq} \|A\|_2 \cdot \|v\|$

$$\|Av\| \leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \|v\|$$

If $\|v\| = 0$ this holds since $0 = \|A\|_2 \cdot 0$. Otherwise we divide by $\|v\|$ to obtain:

$$\frac{\|Av\|}{\|v\|} \leq \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

This clearly holds.

Secondly,

$$\|\Delta x\| = \|A^{-1} \Delta b\| \leq \|A^{-1}\|_2 \|\Delta b\| \frac{\|b\|}{\|b\|}$$

$$\begin{aligned}
&= \|A^{-1}\|_2 \|\Delta b\| \frac{\|Ax\|}{\|b\|} \leq \|A^{-1}\|_2 \|\Delta b\| \|A\|_2 \|x\| \frac{1}{\|b\|} \\
&= \kappa(A) \frac{\|\Delta b\|}{\|b\|} \cdot \|x\|.
\end{aligned}$$

From this the proposition directly follows.

- b) An example where the relative noise on the output $\frac{\|\Delta b\|}{\|b\|} \leq 0.001$ and the relative noise on the input $\frac{\|\Delta x\|}{\|x\|} \geq 1$ is given by the following matrix:

$$A = \begin{pmatrix} 10000 & 0 \\ 0 & 1 \end{pmatrix}$$

and the vector $\Delta b = (0, 1)$. In this case the input vector is given by $x = (1, 0)$ and such $\Delta x = (0, 1)$ and thus the relative noise of the input $\frac{\|\Delta x\|}{\|x\|} = 1$ and the output noise is $\frac{1}{10000} \leq 0.001$.