Aalto University

Linear algebra

Exercise sheet 12 / Model solutions

1. Let $A \in \mathbb{R}^{m\times n}$ have singular value decomposition $A = U\Sigma V^T$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal (i.e., $U^TU = I \in \mathbb{R}^{m\times m}$ and $V^TV = I \in \mathbb{R}^{n\times n}$) and $\Sigma \in \mathbb{R}^{m\times n}$ is the diagonal matrix such that $\Sigma_{ii} = \sigma_i$, where σ_i are the singular values of A. The Moore–Penrose pseudoinverse of A is defined as

$$A^\dagger = V\Sigma^\dagger U^T \in \mathbb{R}^{n\times m},$$

where $\Sigma^\dagger \in \mathbb{R}^{n\times m}$ is the diagonal matrix with entries

$$\Sigma^\dagger_{ii} = \begin{cases} \sigma_i^{-1} & \text{when } \sigma_i > 0, \\ 0 & \text{when } \sigma_i = 0. \end{cases}$$

   (a) Show that $P = AA^\dagger \in \mathbb{R}^{m\times m}$ is an orthogonal projection to the subspace $R(A)$.

   (b) Show that $\boldsymbol{x}^\dagger = A^\dagger \boldsymbol{b} \in \mathbb{R}^n$ is a solution to the equation

$$A\boldsymbol{x} = P\boldsymbol{b}. \tag{1}$$

   (c) Show that $\boldsymbol{x}^\dagger$ is a solution to the least-squares problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2.$$

Solution:

   (a) Observe that $P = AA^\dagger = U\Sigma V^T V\Sigma^\dagger U^T = U\Pi U^T$ where $\Pi \in \mathbb{R}^{m\times m}$ is a diagonal matrix such that $\Pi_{ii} = 1$ when $\sigma_i > 0$ and $\Pi_{ii} = 0$ when $\sigma_i = 0$ or $i > n$. Let r be the largest value of i such that $\sigma_i > 0$; then the first r columns of U are orthonormal and they span $R(A)$, since we have $A\boldsymbol{x} = U(\Sigma V^T\boldsymbol{x})$. It follows that $P = U_rU_r^T$, where U_r is the matrix obtained by taking the first r columns of U, and hence an orthonormal projection onto $R(A)$.

   (b) We have $A\boldsymbol{x}^\dagger = AA^\dagger\boldsymbol{b} = P\boldsymbol{b}$.

   (c) Equivalently, $\boldsymbol{x}^\dagger$ solves the equation $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$:

$$A^TA\boldsymbol{x}^\dagger = A^TAA^\dagger\boldsymbol{b} = V\Sigma^TU^TU_rU_r^T\boldsymbol{b} = V\Sigma^TU^T\boldsymbol{b} = A^T\boldsymbol{b}.$$

2. Let $A \in \mathbb{R}^{n\times n}$ have singular values $\sigma_1, \ldots, \sigma_n$ such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. Show that:

   (a) $\|A\|_2 = \sigma_1$,

(b) $\|A^{-1}\|_2 = \frac{1}{\sigma_n}$ (when A is invertible),

(c) $\kappa_2(A) = \frac{\sigma_1}{\sigma_n}$ (when A is invertible).

Solution: Let $A = USV^T$ be a singular value decomposition of A.

(a) For any $\boldsymbol{x} \in \mathbb{R}^n$ we have

$$\|A\boldsymbol{x}\|_2^2 = \boldsymbol{x}^T A^T A\boldsymbol{x} = \boldsymbol{x}^T VSU^T USV^T \boldsymbol{x} = \boldsymbol{x}^T VS^2V^T \boldsymbol{x},$$

where we used that $U^TU = I$. Let us denote $\boldsymbol{y} = V^T\boldsymbol{x}$. Since V is unitary, $\|\boldsymbol{y}\|_2 = \sqrt{\boldsymbol{y}^T\boldsymbol{y}} = \sqrt{\boldsymbol{x}^T VV^T \boldsymbol{x}} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}} = \|\boldsymbol{x}\|_2$. We get

$$\frac{\|A\boldsymbol{x}\|_2^2}{\|\boldsymbol{x}\|_2^2} = \frac{\boldsymbol{y}^T S^2 \boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}} = \frac{\sum_{i=1}^n y_i^2 \sigma_i^2}{\sum_{i=1}^n y_i^2}.$$

This attains the maximum value when $y_1 \neq 0$ and $y_i = 0$ for all $i \in \{2, \dots, n\}$, since

$$\frac{\sum_{i=1}^n y_i^2 \sigma_i^2}{\sum_{i=1}^n y_i^2} \leq \frac{\sum_{i=1}^n y_i^2 \sigma_1^2}{\sum_{i=1}^n y_i^2} = \sigma_1^2$$

and we can select such a $\boldsymbol{y}$ because V is unitary. And like this we get $\|A\|_2^2 = \sigma_1^2$.

(b) Let A be invertible, and $\sigma_i > 0$ for all $i = 1, \dots, n$. Then we have

$$A^{-1} = V^{-T} S^{-1} U^{-1} = V \operatorname{diag}(\sigma_1^{-1}, \dots, \sigma_n^{-1}) U^T.$$

By part (a) we get

$$\|A^{-1}\|_2 = \sigma_1(A^{-1}) = \max\{\sigma_1^{-1}, \dots, \sigma_n^{-1}\} = \sigma_n(A)^{-1}.$$

(c) By definition and bty the previous parts we get $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \sigma_1\sigma_n^{-1}$.