

1. Least Squares and Gram-Schmidt

Consider the least squares problem

$$\vec{x}^* = \operatorname{argmin}_{\vec{x} \in \mathbb{R}^n} \left\| A\vec{x} - \vec{b} \right\|_2^2 \quad (1)$$

where $A \in \mathbb{R}^{m \times n}$, $\vec{b} \in \mathbb{R}^m$ and assume A is full column rank. One way to solve this least-squares problem is to use Gram-Schmidt Orthonormalization (GSO). Using GSO, the matrix A can be written as,

$$A = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad (2)$$

where Q is an orthonormal matrix and R is an upper-triangular matrix. The columns of Q_1 form an orthonormal basis for the range space $\mathcal{R}(A)$ and columns of Q_2 form an orthonormal basis for the range space $\mathcal{R}(A)^\perp$. Moreover, R_1 is upper triangular and invertible.

(a) Show that the squared norm of the residual is given by

$$\|\vec{r}\|_2^2 := \left\| \vec{b} - A\vec{x} \right\|_2^2 = \left\| Q_1^\top \vec{b} - R_1 \vec{x} \right\|_2^2 + \left\| Q_2^\top \vec{b} \right\|_2^2. \quad (3)$$

(b) Find \vec{x}^* such that the squared norm of the residual in Equation (3) is minimized. Your expression for \vec{x}^* should only use some or all of the following terms: Q_1 , Q_2 , R_1 , \vec{b} .

(c) Check if the expression for \vec{x}^* obtained in the previous part is equivalent to the one obtained by the formula, $\vec{x}^* = (A^\top A)^{-1} A^\top \vec{b}$.

2. Eigenvalues

- (a) Let $A \in \mathbb{R}^{n \times n}$ be a matrix with eigenvalues and corresponding eigenvectors $\lambda_1, \dots, \lambda_n$ and $\vec{v}_1, \dots, \vec{v}_n$ respectively. Now consider $B = A + cI_n$ where $c \in \mathbb{R}$ and I_n is the $n \times n$ identity matrix. What are the eigenvalues and eigenvectors of B in terms of c and λ_i, \vec{v}_i for $i = 1, \dots, n$?

- (b) Let Q be an orthonormal matrix, i.e., $Q^\top Q = I$ of size $n \times n$. Let $\lambda \in \mathbb{R}$ be a scalar, $\vec{v} \in \mathbb{R}^n$ be a vector, and $A \in \mathbb{R}^{n \times n}$ be a matrix. Prove that if we have

$$A\vec{v} = \lambda\vec{v}, \quad (4)$$

i.e., \vec{v}, λ is an eigenpair of A , then we have

$$(QAQ^\top)(Q\vec{v}) = \lambda(Q\vec{v}). \quad (5)$$

- (c) Let A be a $d \times n$ matrix. Prove that the non-zero eigenvalues of AA^\top are the same as the non-zero eigenvalues of $A^\top A$.

- (d) Given a matrix $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ find its eigenvalues and eigenvectors without using the characteristic polynomial.

HINT: Use the fact that eigenvectors with eigenvalue 0 span the null space. Also notice that the eigenvectors of a symmetric matrix corresponding to different eigenvalues are orthogonal to each other.

The reason to not use the characteristic polynomial method for certain matrices is that computing determinants can be expensive in general. But when we have matrices with more structure, we can simplify the problem.

3. Symmetric Matrices

- (a) Show that any symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive semidefinite if and only if there exists a symmetric matrix $C \in \mathbb{R}^{n \times n}$ such that $A = C^\top C$.

(b) Draw the region $\left\{ \vec{x} \in \mathbb{R}^2 \mid \vec{x}^\top \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \vec{x} \leq 1 \right\}$.

(c) Draw the region $\left\{ \vec{x} \in \mathbb{R}^2 \mid \vec{x}^\top \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \vec{x} \leq 1 \right\}$.

- (d) Why is the region in part (b) bounded, whereas the region in part (c) is unbounded?