Quiz01

1. Least Squares and Gram-Schmidt

Consider the least squares problem

$$\vec{x}^\star = \underset{\vec{x}\in\mathbb{R}^n}{\operatorname{argmin}} \left\|A\vec{x} - \vec{b}\right\|_2^2 \tag{1}$$

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,

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

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

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

Solution: We have,

$$\|\vec{r}\|_2^2 := \left\|\vec{b} - A\vec{x}\right\|_2^2 \tag{4}$$

$$= \left\|\vec{b} - Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}\vec{x}\right\|_2^2. \tag{5}$$

Since multiplying by an orthonormal matrix does not change the ℓ_2-norm of a vector we can multiply by $Q^\top$ to get,

$$\|\vec{r}\|_2^2 = \left\|Q^\top \left(\vec{b} - Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}\vec{x}\right)\right\|_2^2 \tag{6}$$

$$= \left\|\begin{bmatrix} Q_1^\top \vec{b} \\ Q_2^\top \vec{b} \end{bmatrix} - \begin{bmatrix} R_1\vec{x} \\ 0 \end{bmatrix}\right\|_2^2 \tag{7}$$

$$= \left\|\begin{bmatrix} Q_1^\top \vec{b} - R_1\vec{x} \\ Q_2^\top \vec{b} \end{bmatrix}\right\|_2^2 \tag{8}$$

$$= \left\|Q_1^\top \vec{b} - R_1\vec{x}\right\|_2^2 + \left\|Q_2^\top \vec{b}\right\|_2^2. \tag{9}$$

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

Solution: We have,

$$\|\vec{r}\|_2^2 = \left\|Q_1^\top \vec{b} - R_1\vec{x}\right\|_2^2 + \left\|Q_2^\top \vec{b}\right\|_2^2. \tag{10}$$

Since we have no control over the term $\left\|Q_2^\top \vec{b}\right\|_2^2$ (i.e., no matter how we change $\vec{x}$, that term stays constant because it doesn't involve $\vec{x}$ at all), it is irrelevant from the perspective of the optimization, and so the optimal $\vec{x}^\star$ is one which minimizes $\left\|Q_1^\top \vec{b} - R_1\vec{x}\right\|_2^2$. This expression is minimized when $Q_1^\top \vec{b} = R_1\vec{x}$, and using the fact that R_1 is invertible we have $\vec{x}^\star = R_1^{-1}Q_1^\top \vec{b}$.

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

Solution: We have $A = QR = Q_1 R_1$ (block multiplication for matrices). Substituting,

$$\vec{x}^\star = (R_1^\top R_1)^{-1} R_1^\top Q_1^\top \vec{b} \tag{11}$$

$$= R_1^{-1} (R_1^\top)^{-1} R_1^\top Q_1^\top \vec{b} \tag{12}$$

$$= R_1^{-1} Q_1^\top \vec{b}. \tag{13}$$

This is the same as we got in the previous part.

