EECS 127/227AT Optimization Models in Engineering
Spring 2020 Discussion 0

1. Least squares and Gram-Schmidt

Consider the least squares problem,

$$\vec{x}^* = \arg\min_{\vec{x}\in\mathbb{R}^n} \left\|A\vec{x} - \vec{b}\right\|_2^2,$$

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 problems is to use Gram-Schmidt Orthonormalization (GSO). Using GSO, the matrix A can be written as,

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

where Q is an orthonormal (orthogonal) matrix and R is an upper-triangular matrix.

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

1. Show that the squared norm of the residual is given by

$$\|\vec{r}\|_2^2 \triangleq \|\vec{b} - A\vec{x}\|_2^2 = \left\|\left(Q_1^\top \vec{b} - R_1\vec{x}\right)\right\|_2^2 + \left\|Q_2^\top \vec{b}\right\|_2^2. \tag{1}$$

Solution: We have,

$$\begin{aligned}\|r\|_2^2 &\triangleq \|\vec{b} - A\vec{x}\|_2^2 \\ &= \left\|\vec{b} - Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}\vec{x}\right\|_2^2.\end{aligned}$$

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

$$\begin{aligned}\|\vec{r}\|_2^2 &= \left\|Q^\top\left(\vec{b} - Q\begin{bmatrix} R_1 \\ 0 \end{bmatrix}\vec{x}\right)\right\|_2^2 \\ &= \left\|\begin{bmatrix} Q_1^\top \vec{b} \\ Q_2^\top \vec{b} \end{bmatrix} - \begin{bmatrix} R_1\vec{x} \\ 0 \end{bmatrix}\right\|_2^2 \\ &= \left\|\begin{bmatrix} Q_1^\top \vec{x} - R_1\vec{x} \\ Q_2^\top \vec{b} \end{bmatrix}\right\|_2^2 \\ &= \left\|\left(Q_1^\top \vec{b} - R_1\vec{x}\right)\right\|_2^2 + \left\|Q_2^\top \vec{b}\right\|_2^2.\end{aligned}$$

2. Find $\vec{x}^*$ such that the squared norm of the residual in Equation (1) is minimized. Your expression for $\vec{x}^*$ should only use some or all of the following terms: Q_1,Q_2,R_1,$\vec{b}$.

   Solution: We have,

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

   Since we have no control over the term $\left\|Q_2^\top \vec{b}\right\|_2^2$, the optimal $\vec{x}^*$ is one which minimizes $\left\|\left(Q_1^\top \vec{b} - R_1\vec{x}\right)\right\|_2^2$ and using the fact that R_1 is invertible we have $\vec{x}^* = R_1^{-1}Q_1^\top \vec{b}$.

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

   Solution: We have $A = QR = Q_1R_1$ (Block multiplication for matrices). Substituting,

$$\vec{x}^* = (R_1^\top R_1)^{-1}R_1^\top Q_1^\top \vec{b}.$$

   Check that $(R_1^\top R_1)^{-1}R_1^\top$ is the inverse of R_1 by both left-multiplying and right-multiplying by R_1. The equivalence between the two forms for $\vec{x}^*$ follows.

2. Eigenvalues

Let $A \in \mathbb{R}^{n\times n}$ have the eigendecomposition $P\Lambda P^{-1}$ where Λ is a diagonal matrix with entries consisting of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Note that this is equivalent to stating that A is diagonalizable via the transformation,

$$P^{-1}AP = \Lambda.$$

1. Show that $A^m = P\Lambda^m P^{-1}$, for integer $m \geq 1$.

   Solution:

$$\begin{aligned} A^m &= (P\Lambda P^{-1})(P\Lambda P^{-1})\ldots(P\Lambda P^{-1}) \quad (m\ times) \\ &= P\Lambda(P^{-1}P)\Lambda(P^{-1}P)\ldots\Lambda(P^{-1}P)\Lambda P^{-1} \\ &= P\Lambda^m P^{-1}. \end{aligned}$$

   The last equality follows from the repeated application of the identity $P^{-1}P = I$.

2. Show that determinant of A is the product of its eigenvalues, i.e.

$$\det(A) = \prod_{i=1}^{n} \lambda_i.$$

   Hint: We have the identity $\det(XY) = \det(X)\det(Y)$.

   Solution:
   Write down eigendecomposition of A and use properties of determinant given in the hint.

$$\begin{aligned} \det(A) &= \det(P\Lambda P^{-1}) \\ &= \det(P)\det(\Lambda)\det(P^{-1}) \\ &= \det(PP^{-1})\det(\Lambda) \\ &= \det(\Lambda) \\ &= \prod_{i=1}^{n} \lambda_i \end{aligned}$$