MAT 653: Statistical Simulation

Instructor: Dr. Wei Li Scribe: Jianchen Wei

Sep 9th, 2021

Solving Linear Least Squares

$Ax = b$, given A and b, to find x.

When there is no solution: $b \notin \text{Col}(A)$

In this case, a natural problem to consider is the linear least square problem:

$$\min_x \|Ax - b\|$$

Our strategy is to find a $\hat{x}$ such that $A\hat{x} = \hat{b}$ is a projection of b on $\text{Col}(A)$. Then $\hat{x}$ is an approximate solution of the original problem $Ax = b$, but only in the sense that Ax is close to b in L_2 norm.

Consider $\hat{b} = A\hat{x}$. Obtain $\hat{b}$ by dropping a perpendicular line from b to $\text{Col}(A)$, that is to say $\langle r, \hat{a} \rangle = 0, \hat{a} \in \text{Col}(A)$, where $r = b - \hat{b}$ is the residual vector. In particular,

$$\begin{aligned}
\langle r, a_i \rangle = 0, i = 1, 2, \cdots, n &\iff A^T r = 0 \\
&\iff A^T(b - A\hat{x}) = 0 \\
&\iff A^T A\hat{x} = A^T b
\end{aligned}$$

Now $\hat{x} = (A^T A)^{-1} A^T b$ if $A^T A$ is invertible, and $\hat{b} = A(A^T A)^{-1} A^T b$.

Note: $A^T A\hat{x} = A^T b$ is called Normal Equation.

Fact 1: Normal equation always has at least one solution.

Fact 2: $A^T A$ is invertible if and only if A has full-column rank.

Fact 3: The solution to the normal equation is not necessary a solution to the original problem $Ax = b$.

Example: $A^A x = A^T b$ has a solution but $Ax = b$ does not.

Let $A = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Then $A^T A = 1, A^T b = 1, \hat{x} = 1$. But $Ax = \begin{bmatrix} 0 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which leads to contradiction. So no solution for $Ax = b$.

Remark: $(A^T A + \lambda I)x = A^T b, \forall \lambda > 0$ is called ridge regression. In this case, we have

$$\hat{x} = (A^T A + \lambda I)^{-1} A^T b \implies \hat{y} = A\hat{x}$$

Application of QR to solve a linear least square problem: $\min_x \|Ax - b\|$.

A columns are linearly independent. $A^T Ax = A^T b$. Decompose matrix A by using QR decomposition:

$$\begin{aligned} A = Q_1 R \implies & R^T Q_1^T Q_1 R x = R^T Q_1^T b \\ \implies & R^T R x = R^T Q_1^T b \\ \implies & (R^T)^{-1} R^T R x = (R^T)^{-1} R^T Q_1^T b \\ \implies & R x = Q_1^T b = y \end{aligned}$$

Algorithm:

1. Compute thin QR: $A = Q_1 R$.
2. Compute the vector: $y = Q_1^T b$.
3. Backsolve the $Rx = y$ to obtain $\hat{x}$.

Note: R command: solve(A,b), qr.solve(A,b)

Eigenvalue decomposition

Square matrix M. $Mx = \lambda x$ for some nonzero vector x and some λ(scalar).

We call λ a eigenvalue of M whose associated eigenvector is x.

Note: A is a square matrix. $\det(A) = \prod_{i=1}^{n} \lambda_i$.

Eigen-decomposition: A is a square and symmetric matrix. then $A = P \Lambda P^T$ where

$$P = [u_1, u_2, \cdots, u_n], u\text{'s are the orthonormal eigenvector of } A$$

$$\Lambda = \{\text{diagonal entries are the eigenvalues of } A\}$$

Note: If this symmetric matrix A is invertible, then $A^{-1} = P \Lambda^{-1} P^T$ where

$$\Lambda^{-1} = \begin{pmatrix} \ddots & 0 & 0 \\ 0 & \frac{1}{\lambda_i} & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

Example: $A = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}$, $Av = \lambda v \implies (A - \lambda I)v = 0$. Here we need a non-zero solution, so

$$\det(A - \lambda I) = \begin{vmatrix} 9 - \lambda & 0 \\ 0 & 4 - \lambda \end{vmatrix} = 0 \implies \lambda_1 = 9, \lambda_2 = 4$$

For $\lambda_1 = 9$, since we know $Av_1 = 9v_1 \implies (A - 9I)v_1 = 0$, then we have

$$\begin{pmatrix} 0 & 0 \\ 0 & -5 \end{pmatrix} v_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

For $\lambda_2 = 4$, since we know $Av_2 = 4v_2 \implies (A - 4I)v_2 = 0$, then we have

$$\begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} v_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Now $A = P\Lambda P^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ where $P = [v_1, v_2]$.