MAT 653: Statistical Simulation

Instructor: Dr. Wei Li

2023-08-29

Back **Updated:** 2023-08-29 Statistical Simulation, Wei Li

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 $P = [v_1, v_2]$.

Singular value decomposition

A, a $m \times n$ matrix, $A^T A$ is a symmetric matrix. Let $\{v_1, \ldots, v_n\}$ consists of all orthonormal eigenvectors of $A^T A$. Let $\lambda_1, ..., \lambda_n$ the associated eigenvalues, $\lambda_1 \geq \lambda_2 \geq \lambda_n \geq 0$.
We have $\|Av_i\|^2 = \lambda_i$. Since

$$A^T A v_i = \lambda_i v_i \Rightarrow v_i^T A^T A v_i = \lambda_i v_i^T v_i \lambda_i$$

Denote $\sigma_i = \sqrt{\lambda_i}$, λ_i is the ith eigenvalue of $A^T A$.

Fact: rank of A is equal to the number of positve singular value of A.

SVD: Let A be a rank r matrix. There exists a matrix $\Sigma_{m\times n}$ with diagonal entries in D are the first r singular value. That is

$$\Sigma_{m\times n} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

The decomposition of A is:

$$A = U_{m\times n} \Sigma_{m\times n} V_{n\times n}^T$$

Where U and V are bothy orthogonal matrices. V consists of all the eigenvectors of $A^T A$. U consists the column u_i as

$$u_i = \frac{Av_i}{\|Av_i\|} = \frac{Av_i}{\sigma_i}, 1 \leq i \leq r$$

Those $\{u_i\}_1^r$ can extend to $\{u_i\}_1^m$ as the orthonormal basis.
Matrix U and V have the property:

$$Col(U) = Col(A), Col(V) = Row(A)$$

Reduced SVD

For those $A_{m\times n}$, $U_{m\times m}$, $V_{n\times n}$, $\Sigma_{m\times n}$ above, we have the partition:

$$U = [U_r, U_{m-r}], V = [V_r, V_{n-r}], \Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

Consequently, the SVD of A can be represented as:

$$A = \begin{bmatrix} U_r & U_{m-r} \end{bmatrix} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r D V_r^T$$

Application

1. Linear Least Squares

We can use SVD to solve the Least Squares Problems as following:

$$
\begin{aligned}
Ax &= b,\\
A^TAx &= A^Tb,\\
(UDV^T)^T(UDV^T)x &= (UDV^T)^Tb,\\
\Rightarrow VD^2Dx &= VDU^Tb,\\
\Rightarrow D^2V^Tx &= DU^Tb,\\
\Rightarrow DV^Tx &= U^Tb.
\end{aligned}
$$

Denote $w = V^Tx$,$y = U^Tb$, we have the following algorithm:

Algorithmn

1. Find the SVD of $A = UDV^T$;

2. Compute $y = U^Tb$;

3. Solve $w^* = D^{-1}y$;

4. $V^Tx = w^* \Rightarrow x = Vw^*$;

The solution of the equation is $x = V_rD^{-1}U_r^Tb$, and we can notice that this SVD method allows A,b to be arbitrary. Notice that $V_rD^{-1}U_r^T$ is the inverse of A. Here we have the concept of generalized inverse.

Moore-Penrose inverse: $A^+ = V_rD^{-1}U_r^T$

2.LS-Problem

In LS Problem

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

,

where A and b are not restricted at all, we have:

$$A^TAx = A^Tb.$$

Let $\mathfrak{L}$ is the set of all the minimizers to the LS problem. We have following facts.

Fact 1 $x^* = A^+b \in \mathfrak{L}$.

Fact 2 $A\tilde{x_1} = A\tilde{x_2}$, for any $\tilde{x_1}, \tilde{x_2} \in \mathfrak{L}$.

Fact 3 For the optimization problem $\min_{x\in\mathfrak{L}} \|x\|$, there is a unique solution $x^* = A^+b$.

Fact 4 We already have the result that if we choose some λ, LS problem will has a unique solution to the "modified" normal equation:

$$(A^TA + \lambda I)x = A^Tb$$

that is,

$$\hat{x} = (A^TA + \lambda I)^{-1}A^Tb$$

In fact, let $\lambda \to 0$, we have

$$(A^TA + \lambda I)^{-1}A^T \to A^+$$

Fact 5 The projection of b on Col(A) is given by $A(A^TA)^{-1}A^Tb$ (assuming A^TA is invertible), here is a more general result

$$\hat{b} = AA^+b = \lim_{\lambda \to 0}[A(A^TA + \lambda I)^{-1}A^T]b$$

Back **Updated:** 2023-08-29 Statistical Simulation, Wei Li