Lecture 2-Brushing Up Matrices

Multiplying and Factoring Matrices

Today's Discussion

- Five important Factorizations
- $A=LU$
- $A=QR$
- $S=Q\boldsymbol{\Lambda}Q^T$
- $A=X\boldsymbol{\Lambda}X^{-1}$
- $A=U\boldsymbol{\Sigma}V^T$

Source: Section I.2 in Linear Algebra and Learning from Data (2019) by **Gilbert Strang**

Multiplying and Factoring Matrices

1. A=LU (elimination- solving linear systems). Matrix *L* is lower triangular, and *U* is upper triangular. but it's about elimination. Solving linear systems. (Discuss later)

2. A=QR-> Least Squares, the big application, the factorization. Letter Q? Orthogonal. Columns are orthogonal. Often orthonormal. Orthogonal means they're perpendicular to each other. And orthonormal means they're unit vectors. Q often represents a matrix with orthonormal columns. So, we say Gram-Schmidt, whose algorithm produces Q and R. Matrix Q is orthonormal columns QQ^T= I

$$Q^TQ = \begin{bmatrix} - & \boldsymbol{q}_1^{\mathrm{T}} & - \\ - & \boldsymbol{q}_2^{\mathrm{T}} & - \\ & & \\ - & \boldsymbol{q}_n^{\mathrm{T}} & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \boldsymbol{q}_1 & \boldsymbol{q}_2 & & \boldsymbol{q}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$$

$$\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_j = \begin{cases} 0 & \text{when } i \neq j \quad (\textbf{\textit{orthogonal}} \text{ vectors}) \\ 1 & \text{when } i = j \quad (\textbf{\textit{unit}} \text{ vectors: } \|\boldsymbol{q}_i\| = 1) \end{cases}$$

$\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_i = \|\boldsymbol{q}_i\|^2 = 1.$ Often Q is rectangular (m>n). Sometimes m=n

When Q is square, Q^TQ=I means that $Q^T=Q^{-1}$: transpose = inverse

Figure Source: Gilbert Strang - Introduction to Linear Algebra

Multiplying and Factoring Matrices

3. $S=Q\Lambda Q^T$. S stands for symmetric i.e., $S=S^T$. A special factorization for symmetric matrices
λ is diagonal eigen value matrices- λ always for eigenvalues
Q has orthonormal eigenvectors (Discuss later)

4. $A=X\Lambda X^{-1}$. A has a set of n independent eigen vectors; $x_1, x_2,.. x_n$. Multiply $A\ X$ column by column to get the columns $\lambda_1 x_1$ to $\lambda_n x_n$. This matrix split into X times Λ

$$A\begin{bmatrix} \boldsymbol{x}_1 & .. & \boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} A\boldsymbol{x}_1 & .. & A\boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{x}_1 & .. & \lambda_n\boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & .. & \boldsymbol{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

This equation gives $AX= \Lambda X$ that implies that $A=X\Lambda X^{-1}$. if we know eigen values and eigen vectors, and matrix A, then we compute powers of A

$$\boxed{\begin{array}{llll} \boldsymbol{\Lambda} & = \text{diagonal eigenvalue matrix} & A & = X\Lambda X^{-1} \\ \boldsymbol{X} & = \text{invertible eigenvector matrix} & A^2 & = (X\Lambda X^{-1})(X\Lambda X^{-1}) = X\Lambda^2 X^{-1} \end{array}}$$

Note: $(X\Lambda X^{-1})(X\Lambda X^{-1})=X(X\Lambda X^{-1}) \times \Lambda\ (X\Lambda X^{-1}) \times X^{-1}(X\Lambda X^{-1})=X\Lambda^2 X^{-1}$ (since $X\Lambda=I$ and $XX^{-1}=I$ and $X^{-1}X=I$)

Figure Source: Linear Algebra and Learning from Data -Gilbert Strang

Multiplying and Factoring Matrices

5. $A=U\Sigma V^T$ the Singular Value Decomposition (SVD) of matrix A (square or not). Singular values $\sigma_1, \sigma_2 \ldots \sigma_r$ in Σ. U and V are orthogonal matrices (Discuss later)

Let us discuss $S=Q\Lambda Q^T$

Symmetric matrix S	$S^{\mathrm{T}} = S$	All $s_{ij} = s_{ji}$
Orthogonal matrix Q	$Q^{\mathrm{T}} = Q^{-1}$	All $q_i \cdot q_j = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$

$$SQ = S\begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} \lambda_1 q_1 & \cdots & \lambda_n q_n \end{bmatrix} = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = Q\Lambda$$

Multiply $SQ= Q\Lambda$ by $Q^{-1} = Q^T$ which gives us $S=Q\Lambda Q^T =$ a symmetric matrix

For Exp: $(Q\Lambda)(Q^T)$= colmns of $Q\Lambda$ x (rows of Q^T) = $\begin{bmatrix} \ \end{bmatrix}[\ ]$ = [Column space] which has rank=1

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}\begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix} \quad \text{which has rank} = 1$$

Figure Source: Linear Algebra and Learning from Data -Gilbert Strang

Multiplying and Factoring Matrices

$(Q\Lambda)(Q^T)$= Sum of S = $\lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \ldots + \lambda_n q_n q_n^T$

$Sq_1 = \lambda_1 q_1 q_1^T q_1 + \lambda_2 q_2 q_2^T q_1 + \ldots + \lambda_n q_n q_n^T q_1 = \lambda_1 q_1 q_1^T q_1$ as q's are orthogonal other terms are zero.

This Implies that $Sq_1 = \lambda_1 q_1$

$q_1 q_1^T = \| q_1 \|^2 \rightarrow$ Q is orthonormal matrix

Back-up Slides

Multiplying and Factoring Matrices

Example:

Let A= $\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$. Find A^2 given that A= $X\boldsymbol{\Lambda}X^{-1}$ where X=$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, $\boldsymbol{\Lambda}$=$\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}$ and X^{-1} = $\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$

$A^2 = (X\boldsymbol{\Lambda}X^{-1})\ (X\boldsymbol{\Lambda}X^{-1}) = X\boldsymbol{\Lambda}\ (XX^{-1})\ \boldsymbol{\Lambda}X^{-1} = X\boldsymbol{\Lambda}^2X^{-1}$

In general $A^k = X\boldsymbol{\Lambda}^kX^{-1}$ = $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 4^k \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$ = $\begin{bmatrix} 2.5^k - 4^k & -5^k + 4^k \\ 2.5^k - 2.4^k & -5^k + 2.4^k \end{bmatrix}$

A square matrix A is said to be diagonalizable if A is similar to a diagonal matrix, i.e., if A = $X\boldsymbol{\Lambda}X^{-1}$ where X is invertible and $\boldsymbol{\Lambda}$ is a diagonal matrix

When is A diagonalizable? (The answer lies in examining the eigenvalues and eigenvectors of A)

Note that $\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ = 5 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$= 4 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Altogether, $\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ = $\begin{bmatrix} 5 & 4 \\ 5 & 8 \end{bmatrix}$ = $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} - & 0 \\ 0 & - \end{bmatrix}$

Equivalently, $\begin{bmatrix} 6 & -1 \\ 2 & 3 \end{bmatrix}$ = $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$