Lecture 3-Brushing Up Matrices

Multiplying and Factoring Matrices (Contd.)

Multiplying and Factoring Matrices

- Today's Discussion
- A=LU
- A=U$\boldsymbol{\Sigma}$V^T

Source: Section I.4 and I.8 in Linear Algebra and Learning from Data (2019) by **Gilbert Strang**

Multiplying and Factoring Matrices

- Let’s talk about elimination i.e., A=LU. It is solving like Ax=b which are row operations
- All those row operations that is expressed by L times U
- And it factors into lower triangular (L) times upper triangular (U)
- Take A (2x2) matrix, A=$\begin{bmatrix} 2 & 3 \\ 4 & 7 \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{2} & 3 \\ 0 & \mathbf{1} \end{bmatrix}$
- Multiplying row 1 by 2 and subtracting row 1 from row 2
- Upper triangular with the pivots **2** and **1** on the diagonal

Multiplying and Factoring Matrices

- Matrix A is L times U. L is the lower triangular matrix
- Take L= $\begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix}$
- It has there the number that you used here. We multiplied 2 of the first row and subtracted it from 2nd row in A. So, we need a multiplier i.e., 2 there. So,
- L= $\begin{bmatrix} \mathbf{1} & 0 \\ 2 & \mathbf{1} \end{bmatrix}$
- Then it shows
- A= $\begin{bmatrix} \mathbf{1} & 0 \\ 2 & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{2} & 3 \\ 0 & \mathbf{1} \end{bmatrix}$ =LU

Multiplying and Factoring Matrices

Parallel way to think of this 2 by 2 matrix A= $\begin{bmatrix} 2 & 3 \\ 4 & 7 \end{bmatrix}$ = $\begin{bmatrix} 2 & 3 \\ 4 & - \end{bmatrix}$ + $\begin{bmatrix} 0 & 0 \\ 0 & - \end{bmatrix}$ $\rightarrow$ $\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$ + $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ =L+U

Split it into first row and column in one piece

Elimination has taken the original matrix. It's split and these are both rank 1

So, L+U= [col l_1] [row u_1^T] + [col l_2] [row u_2^T]

Our idea is that it gives the breakdown. And this is, by column times row rule, that's LU

$$A = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} = \text{(col1)(row1)} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & A22 & A23 \\ 0 & A32 & A33 \end{bmatrix} = \text{(col1)(row1)} + \text{(col2)(row2)} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A33 \end{bmatrix}$$

Start with A, break it up into LU, where LU is the first column times row. And then the next pieces are the rest of the matrix

Second part, second column times the second row, maybe divide by the pivot to make it correct. And then A3 would be the rest of that matrix

Multiplying and Factoring Matrices

Singular Value Decomposition (SVD) $A=U\boldsymbol{\Sigma}V^T$
Compare with $S=Q\boldsymbol{\Lambda}Q^T$

For a rectangular matrix, the idea of eigenvalues is snapped because if Ax in n dimensions, output will come in m dimensions. $Ax = \boldsymbol{\lambda}x$ is not even possible if A is rectangular

A (rectangular matrix)=$U\boldsymbol{\Sigma}V^T$ ➞ U and V are singular vectors and $\boldsymbol{\Sigma}$ is singular values matrix of rank r, i.e.,

$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\sigma}_1 & & \\ & \boldsymbol{\sigma}_2 & \\ & & \boldsymbol{\sigma}_r \end{bmatrix}$ and other entries are zeros

A^TA =[mxn rows][nxm columns]=(nxn) matrix= Symmetric matrix. Since eigen values are ≥ 0, A^TA is positive definite (p.d) matrix

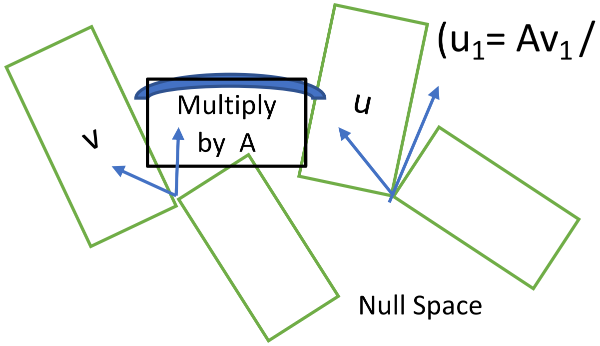
Multiplying and Factoring Matrices

Matrix can be factored $A^TA = V\Lambda V^T$ where V is orthogonal eigen vector matrix, V^T is its transpose and eigen value $\Lambda \geq 0$, positive definite

AA^T = (mxm) large matrix with same eigen values $\Lambda = U\Lambda U^T$, U is eigen vectors. Use U's and V's. They are same.

Matrix A is rectangular and hasn't eigenvectors. Av is sigma, singular value ($\boldsymbol{\sigma}$) times u. That's the first entry and the second entry and the rth entry. Stop at r, that is the rank. So,

$(u_1 = Av_1 / \boldsymbol{\sigma}_1)$



$$\left.\begin{array}{l} Av_1 = \boldsymbol{\sigma}_1 u_1 \\ Av_2 = \boldsymbol{\sigma}_2 u_2 \\ . \\ . \\ Av_r = \boldsymbol{\sigma}_r u_r \end{array}\right\} \qquad A\left[v_1, v_2, \dots v_r\right] = \left[u_1, u_2, \dots u_r\right]\begin{pmatrix} \boldsymbol{\sigma}_1 & & \\ & \boldsymbol{\sigma}_2 & \\ & & \boldsymbol{\sigma}_r \end{pmatrix}$$

$$\Rightarrow AV = \boldsymbol{\Sigma}U \Rightarrow A = U\boldsymbol{\Sigma}V^T$$

$$\left.\begin{array}{l} Av_{r+1} = 0. \\ .. \\ Av_n = 0 \end{array}\right\} \text{Null Space}$$

$A^TA = V\boldsymbol{\Sigma}^TU^TU\boldsymbol{\Sigma}V^T = V(\boldsymbol{\Sigma}^T\boldsymbol{\Sigma})V^T$ where, **V=evector of A^TA** and **$\boldsymbol{\sigma}^2$ evalues of A^TA** and A^TA is symmetric, positive definite

Multiplying and Factoring Matrices

Using A^TA expression in getting u's

$Av_1 / \boldsymbol{\sigma}_1 = u_1 \ldots \ldots Av_r / \boldsymbol{\sigma}_r = u_r$

We've chosen v's and $\boldsymbol{\sigma}$'s. For A^TA, eigenvectors are v's and eigenvalues are sigma squared. Now we want u

$u_1^T u_2 = 0$ (u_1 and u_2 are orthogonal)

$u_1^T u_2 \Rightarrow (Av_1 / \boldsymbol{\sigma}_1)^T (Av_2 / \boldsymbol{\sigma}_2) \Rightarrow (v_1 A^T A v_2 / \boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2) \Rightarrow v_1^T \boldsymbol{\sigma}_2^2 v_2 / \boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2$ (v_1 and v_2 are orthogonal and v_2 evector of A^TA)
$\Rightarrow (\boldsymbol{\sigma}_2 / \boldsymbol{\sigma}_1)\, v_1^T v_2 = 0$

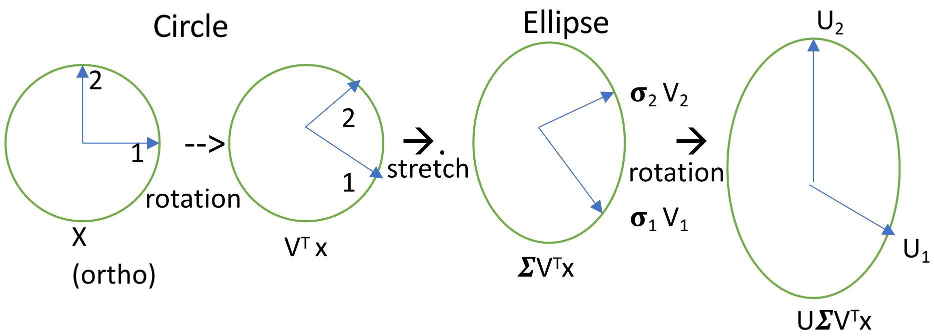
If we have a matrix A, say 5,000 by 10,000, why is it a mistake to use A^TA in the computation? It's very big and very expensive.

Actual computational methods are quite different

Because A^TA, it's symmetric, positive definite, we make this proof

Multiplying and Factoring Matrices

We take $A=U\boldsymbol{\Sigma}V^T$ (U and V ortho. Matrices) Then $Ax=U\boldsymbol{\Sigma}V^Tx$



This SVD is a linear transformation and every linear transformation, i.e., every matrix multiplication factors into a rotation times a stretch times a different rotation.

when u be the same as a v?

A square. Because A is a square matrix, and this is the same as $Q\Lambda Q^T$. If they're the same, U's are the same as the V's when the matrix is symmetric
And these need to be positive definite as $\sigma_1 > \sigma_2 > \ldots > \sigma_r > 0$

$A= U\Sigma V^T$ So $A=Q\Lambda Q^T$

Why is that?
Because σ's are ≥ 0. and a positive definite symmetric matrix, S is the same as the A. Q is the U, the Q^T is the V^T, then λ is the σ

Multiplying and Factoring Matrices

$$\underset{n\times m}{A^T}\ \underset{m\times n}{A} = \underset{n\times n}{V\, \Lambda\, V^T} \quad \text{(V is orthogonal and } \Lambda \geq 0)$$

$$AA^T = \underset{n\times n}{U\, \Lambda\, U^T}$$

$$A = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix}$$

$$\sigma_1 \leq \lambda_1 \leq \lambda_2 \leq \sigma_2$$

$$\sigma_1\sigma_2 = 15 = \sigma_1\sigma_2$$

$$A = \begin{array}{c} (m\times r) \\ \begin{pmatrix} u_{1.} & u_{2.} & u_r \end{pmatrix} \end{array} \begin{array}{c} (r\times r) \\ \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_r \end{pmatrix} \end{array} \begin{array}{c} (r\times n) \\ \begin{pmatrix} v_1^T \\ v_2^T \\ v_r^T \end{pmatrix} \end{array}$$

$$AA^T = \begin{array}{c} (n\times m) \\ \begin{pmatrix} u_{1.} & u_{2.} & u_m \end{pmatrix} \end{array} \begin{array}{c} (m\times n) \\ \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_r \\ 0 & 0 \ldots & 0 \end{pmatrix} \end{array} \begin{array}{c} (n\times n) \\ \begin{pmatrix} v_1^T \\ v_2^T \\ v_n^T \end{pmatrix} \end{array}$$

$$= u_1\sigma_1 v_1^T = \text{rank 1 matrix}$$

Multiplying and Factoring Matrices

$A = U\Sigma V^T = SQ$

From SVD,

$(U\Sigma V^T)\,(UV^T) = SQ$ ($U\Sigma V^T \rightarrow$ symmetric and $(UV^T) \rightarrow$ orthogonal)

Key fact-- if we have a big matrix of data, A, and if we want to pull out of that matrix the important part, what data science must be doing

Out of a big matrix, some part of it is noise, some part of it is signal

Signal is most important part of the matrix

Multiplying and Factoring Matrices

Let us look at $U\boldsymbol{\Sigma}V^T$

It's a rank one

Simple matrix building block is a rank one matrix, a something (U), something transpose (V^T)

That is, $u_1\boldsymbol{\sigma}_1 v_1^T$ = rank 1 matrix

Back-up Slide

Multiplying and Factoring Matrices

$AA^T = U\boldsymbol{\Sigma}V^TV\boldsymbol{\Sigma}^TV^T = U(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T)U^T$ (as $V^TV=I$) U is evectors of AA^T
A^TA and AA^T are two different. So each has its own eigen vectors and we use both. It's just perfect as far as it goes, but it hasn't gone to the end yet because we could have double eigenvalues and triple eigenvalues, and all those horrible possibilities.
Suppose a matrix has a symmetric matrix, has a double eigenvalue. Take an example

$$S = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 5 \end{pmatrix} \begin{pmatrix} X \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ 0 \end{pmatrix}$$

What's the deal with eigenvectors for that matrix 1, 1, 5? So 5 has got an eigenvector. You can see what it is, 0, 0, 1 eigenvectors that go with $\lambda=1$ for that matrix? What would be eigenvectors for a $\lambda=1$? There was a whole plane of them. Any vector of the form x, y, 0, i.e., the whole plane of eigenvectors

And we need to pick two that are orthogonal, which we can do. And then they have to be-- in the SVD those two orthogonal vectors have to go to two orthogonal vectors. We conclude that the V's the singular vectors should be eigenvalues