Math 415 - Lecture 37

Singular Value Decomposition

Friday December 4th 2015

Textbook reading: Chapter 6.3

Suggested practice exercises: Chapter 6.3, # 1, 2, 3, 5, 8, 9, 15

Strang lecture: Lecture 29: Singular Value Decomposition

1 Review

- Spectral theorem: If A is an $n \times n$ symmetric matrix, then it has an orthonormal basis of eigenvectors $\mathbf{v}_1 \dots \mathbf{v}_n$, and all eigenvalues are real.
- We can write

$$A = \underbrace{\begin{bmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{bmatrix}}_{\text{matrix of eigenvectors}} \underbrace{\begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & 0 & & \lambda_n \end{bmatrix}}_{\text{matrix of eigenvalues}} \begin{bmatrix} - & \mathbf{v}_1^T & - \\ & \vdots & \\ - & \mathbf{v}_n^T & - \end{bmatrix}$$

- Today: There is a similar decomposition for any $m \times n$ matrix A.
  - Doesn't even have to be square!
  - The price we pay: different bases on the left and right sides.

1.1 What is SVD?

Motto
In Linear Algebra everything is a matrix factorization.

The complicated story with orthonormal basis and singular values for A gives a factorization, called Singular Value Decomposition:

$$A = \underbrace{\begin{bmatrix} | & & | \\ \mathbf{u}_1 & \dots & \mathbf{u}_m \\ | & & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \\ 0 & \sigma_2 & \\ & & \ddots \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} - & \mathbf{v}_1^T & - \\ & \vdots & \\ - & \mathbf{v}_n^T & - \end{bmatrix}}_{V^T}$$

- $A = U\Sigma V^T$
- U, V are orthogonal
- Columns of U are an orthonormal basis for $\mathbb{R}^m$. U is $m \times m$
- Rows of V are an orthonormal basis for $\mathbb{R}^n$. V is $n \times n$
- Σ is rectangular $m \times n$ and diagonal, the r non zero diagonal entries are called **singular values**, they are positive

1.2 How to Compute SVD

Here is a recipe for computing SVD:

Compute $A^T A$. This is a symmetric matrix!! (Why?)

Make V:
- Find orthonormal eigenvectors $\mathbf{v_1}, \dots, \mathbf{v_n}$ of $A^T A$. (Why can we do this?)
- Of $\mathbf{v_1}, \dots, \mathbf{v_n}$, make the first ones, $\mathbf{v_1}, \dots, \mathbf{v_r}$, be ones with **nonzero** eigenvalues $\lambda_1 \geq \dots \geq \lambda_r > 0$.
- Magic: $\lambda_1, \dots, \lambda_r$ always positive!
- Put $\mathbf{v}_1, \dots, \mathbf{v}_n$ into matrix V, those with non-zero eigenvectors come first!

Make Σ: Set $\sigma_i = \sqrt{\lambda_i}$ for $i = 1 \dots r$ and $\sigma_{r+1} = \cdots = \sigma_n = 0$. Put these into diagonal of **rectangular** $m \times n$ matrix Σ.

Make U:
- Set $\mathbf{u_1} = \frac{1}{\sigma_1} A\mathbf{v_1}$, ..., $\mathbf{u_r} = \frac{1}{\sigma_r} A\mathbf{v_r}$.
- Extend $\mathbf{u_1}, \dots, \mathbf{u_r}$ to an orthonormal basis $\mathbf{u_1}, \dots, \mathbf{u_m}$ for $\mathbb{R}^m$.
- Put $\mathbf{u_1}, \dots, \mathbf{u_m}$ into matrix U.

Now you have $A = U\Sigma V^T$!

Example 1. Compute the SVD of

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Solution.

Example 2. Compute the SVD of

$$A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Solution.

A matrix might not be diagonalizable:

- If A is rectangular, it does not even have eigenvalues.

But A will always have an SVD! This comes at a cost:

- The SVD is not unique.
- The singular values σ_i are not eigenvalues.

Note the difference: for $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ the eigenvalues are $\lambda = i, -i$ but the singular values are $\sigma = 1, 1$.

1.3 Approximation

It turns out we can write A as a sum:

$$A = \underbrace{\begin{bmatrix} | & & | \\ \mathbf{u}_1 & \dots & \mathbf{u}_m \\ | & & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \\ 0 & \sigma_2 & \\ & & \ddots \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} - & \mathbf{v}_1^T & - \\ & \vdots & \\ - & \mathbf{v}_n^T & - \end{bmatrix}}_{V^T}$$

$$= \sigma_1 \mathbf{u_1} \mathbf{v_1}^T + \sigma_2 \mathbf{u_2} \mathbf{v_2}^T + \cdots + \sigma_r \mathbf{u_r} \mathbf{v_r}^T$$

Idea. We can get a good approximation to A by taking the entries of the sum with the largest singular values! We'll see this when we talk about image compression later.

Example 3. If $\mathbf{u}, \mathbf{v}$ are non-zero, then the matrix $\mathbf{u}\mathbf{v}^T$ has rank 1. Why?

Solution.

Example 4. Use

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

to write $\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ as a sum of rank 1 matrices.

Solution.

1.4 SVD and the Four Fundamental Subspaces

The SVD of A gives orthonormal bases for all four fundamental subspaces of A.
Given $\{\mathbf{u_1}, \ldots, \mathbf{u_m}\}$ and $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$,

- $Col(A^T) = \text{Span}\{\mathbf{v_1}, \ldots, \mathbf{v_r}\}$
- $Nul(A) = \text{Span}\{\mathbf{v_{r+1}}, \ldots, \mathbf{v_n}\}$
- $Col(A) = \text{Span}\{\mathbf{u_1}, \ldots, \mathbf{u_r}\}$
- $Nul(A^T) = \text{Span}\{\mathbf{u_{r+1}}, \ldots, \mathbf{u_m}\}$

1.5 Practice Questions

Example 5. Suppose A is an invertible square matrix. Find a singular value decomposition of A^{-1}.

Solution.

Example 6. If A is a square matrix, then $|\det(A)|$ is the product of the singular values of A. Why?

Solution.

Example 7. Find the singular value decomposition of $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$.

Solution.