SINGULAR VALUE DECOMPOSITION

Notes for Math 54, UC Berkeley

Let A be an $m \times n$ matrix. We discuss in these notes how to transform the perhaps complicated A into a simpler form, by multiplying it on the left and right by appropriate orthogonal matrices. This is important for many interesting applications.

LEMMA 1. *The matrix*

$$S = A^T A$$

is a ***symmetric*** $n \times n$ *matrix.*

Proof. We recall the matrix formula $(BC)^T = C^T B^T$, which implies that

$$S^T = (A^T A)^T = A^T (A^T)^T = A^T A = S.$$

The transpose A^T is an $n \times m$ matrix and thus S is $n \times n$. □

Since S is symmetric, it has real eigenvalues $\lambda_1, \dots, \lambda_n$ and corresponding eigenvectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ so that

$$A^T A \mathbf{v}_j = S \mathbf{v}_j = \lambda_j \mathbf{v}_j \qquad (j = 1, \dots, n) \tag{1}$$

and

$$\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \text{ is an orthonormal basis of } \mathbb{R}^n.$$

LEMMA 2. (i) *The following identities hold:*

$$A\mathbf{v}_i \cdot A\mathbf{v}_j = \lambda_j \delta_{ij} \qquad (i, j = 1, \dots, n), \tag{2}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } \; i = j \\ 0 & \text{if } \; i \neq j. \end{cases}$$

(ii) *Furthermore, the eigenvalues of $S = A^T A$ are nonnegative:*

$$\lambda_j \geq 0 \qquad (j = 1, \dots, n).$$

Proof. We use (1) to calculate that

$$A\mathbf{v}_i \cdot A\mathbf{v}_j = (A\mathbf{v}_i)^T A\mathbf{v}_j = \mathbf{v}_i^T A^T A\mathbf{v}_j = \lambda_j \mathbf{v}_i^T \mathbf{v}_j = \lambda_j \mathbf{v}_i \cdot \mathbf{v}_j = \lambda_j \delta_{ij},$$

since $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is orthonormal. In particular, $\lambda_j = ||A\mathbf{v}_j||^2 \geq 0$. □

Let us now reorder, if necessary, the eigenvalues so that

$$\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0.$$

DEFINITION. *The **singular values** of A are the numbers*

$$\boxed{\sigma_j = \sqrt{\lambda_j}} \qquad (j = 1, \dots, n).$$

Then

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0, \tag{3}$$

and formula (2) implies

$$||A\mathbf{v}_j|| = \sigma_j \qquad (j = 1, \dots, n). \tag{4}$$

DEFINITION. *We write*

$$\boxed{\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i} \qquad (i = 1, \dots, r).$$

It follows from (2) and (4) that $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ is orthonormal in $\mathbb{R}^m$, and thus

$$0 \leq r \leq \min\{n, m\}.$$

We can now use the Gram-Schmidt process to find further vectors $\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}$ so that

$$\{\mathbf{u}_1, \dots, \mathbf{u}_m\} \text{ is an orthonormal basis of } \mathbb{R}^m.$$

The key point is that **we can use the orthonomal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ of $\mathbb{R}^m$ and the orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of $\mathbb{R}^n$ to convert our matrix A into a simpler form**. Here is how to do it:

NOTATION. *Introduce the $m \times m$ orthogonal matrix*

$$U = (\mathbf{u}_1|\mathbf{u}_2|\dots|\mathbf{u}_m),$$

whose i^{th} column is $\mathbf{u}_i$ $(i = 1, \dots, m)$. Likewise, introduce the $n \times n$ orthogonal matrix

$$V = (\mathbf{v}_1|\mathbf{v}_2|\dots|\mathbf{v}_n).$$

Then

$$UU^T = U^TU = I, \; VV^T = V^TV = I. \tag{5}$$

THEOREM 1. *We have*

$$U^TAV = \left(\begin{array}{cccc|c} \sigma_1 & 0 & \dots & 0 & \\ 0 & \sigma_2 & \dots & 0 & \\ \vdots & \vdots & \ddots & \vdots & O \\ 0 & 0 & \dots & \sigma_r & \\ \hline & & O & & O \end{array}\right). \tag{6}$$

REMARK. *Thus if we write Σ for the $m \times n$ matrix on the right hand side of* (6)*, we obtain using* (5) *the* ***singular value decomposition (SVD)***

$$\boxed{A = U\Sigma V^T} \tag{7}$$

of our matrix A.

This is similar to the familiar orthogonal diagonalization formula for a symmetric $n \times n$ matrix, but ***in*** (6) ***and*** (7) ***the matrix*** A ***need not be symmetric nor square****.* □

Proof. Since

$$AV = A(\mathbf{v}_1|\mathbf{v}_2|\dots|\mathbf{v}_n) = (A\mathbf{v}_1|A\mathbf{v}_2|\dots|A\mathbf{v}_n),$$

it follows that

$$U^TAV = \begin{pmatrix} \mathbf{u}_1 \cdot A\mathbf{v}_1 & \mathbf{u}_1 \cdot A\mathbf{v}_2 & \dots & \mathbf{u}_1 \cdot A\mathbf{v}_n \\ \mathbf{u}_2 \cdot A\mathbf{v}_1 & \mathbf{u}_2 \cdot A\mathbf{v}_2 & \dots & \mathbf{u}_2 \cdot A\mathbf{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_m \cdot A\mathbf{v}_1 & \mathbf{u}_m \cdot A\mathbf{v}_2 & \dots & \mathbf{u}_m \cdot A\mathbf{v}_n \end{pmatrix}. \tag{8}$$

Now if $j \in \{r+1, \ldots, n\}$, then $A\mathbf{v}_j = 0$. If $j \in \{1, \ldots, r\}$ and $i \in \{r+1, \ldots, m\}$, then

$$\mathbf{u}_i \cdot A\mathbf{v}_j = \sigma_j \mathbf{u}_i \cdot \mathbf{u}_j = 0.$$

Finally, if $i, j \in \{1, \ldots, r\}$, then

$$\mathbf{u}_i \cdot A\mathbf{v}_j = \frac{1}{\sigma_i} A\mathbf{v}_i \cdot A\mathbf{v}_j = \frac{\lambda_i}{\sigma_i} \mathbf{v}_i \cdot \mathbf{v}_j = \sigma_i \delta_{ij}.$$

Using these formulas in (8) gives (6). □

SUMMARY: HOW TO FIND THE SVD

1. Diagonalize $S = A^T A$, to find an orthonormal basis of $\mathbb{R}^n$ of eigenvectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.
2. Reorder the eigenvalues of S so that $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$.
3. Let
$$\sigma_j = \lambda_j^{\frac{1}{2}} \qquad (j = 1, \ldots, n);$$
then
$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$
4. Define
$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i \qquad (i = 1, \ldots, r).$$
5. Extend $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ to an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ of $\mathbb{R}^m$.
6. Write U, V and Σ, as above; then $A = U\Sigma V^T$ is the corresponding singular value decomposition of the matrix A.

EXAMPLE. Find the SVD for the non-symmetric matrix

$$A = \begin{pmatrix} -4 & 6 \\ 3 & 8 \end{pmatrix}.$$

We compute

$$S = A^T A = 25 \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

The eigenvalues of S are $\lambda_1 = 100,\ \lambda_2 = 25$, with corresponding orthonormal eigenvectors

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Therefore

$$\sigma_1 = 10,\ \sigma_2 = 5$$

and

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix},\ \mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{5}\begin{pmatrix} -4 \\ 3 \end{pmatrix}.$$

So

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ U = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix},\ \Sigma = \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}.$$

We check that U, V are orthogonal matrices, and

$$U\Sigma V^T = \frac{1}{5}\begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -4 & 6 \\ 3 & 8 \end{pmatrix} = A. \quad \square$$