The Singular Value Decomposition

Jeremy Teitelbaum

The singular value decomposition (SVD)

The SVD is a way to study rectangular matrices using tools that come from our work with symmetric matrices.

It doesn't make direct sense to diagonalize a rectangular matrix, but in some sense the SVD is the closest we can come.

It is a widely used result in applied mathematics.

Singular Values

Let A by an $m \times n$ matrix. The *singular values* σ_i of A are the (positive) square roots of the eigenvalues of the $n \times n$ symmetric matrix $A^T A$

$$\sigma_i = \sqrt{\lambda_i}$$

Remember that, by the spectral theorem, $A^T A$ has real, nonnegative eigenvalues, so these square roots make sense.

We arrange the singular values in decreasing order so that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

Singular values

If $v_1, \dots, v_n$ are the unit eigenvectors of $A^T A$, then

$$\|Av_i\|^2 = (Av_i) \cdot (Av_i) = v_i^T A^T A v_i = \lambda_i \|v_i\|^2$$

so the singular values σ_i measure the amount that A "stretches" v_i.

Nonzero singular values give rank

Some of the singular values σ_i of A and corresponding eigenvalues λ_i of $A^T A$ could be zero.

If λ_k is zero, then

$$Av_k \cdot Av_k = v_k^T A^T A v_k = \lambda_k (v_k \cdot v_k) = 0$$

so $Av_k = 0$.

Suppose that the first r of them are non zero. Then, if v_i are the corresponding eigenvectors of $A^T A$, the vectors

$$Av_1, \ldots, Av_r$$

form an orthogonal basis for the column space $\mathrm{Col}(A)$, and A has rank r.

Nonzero singular values give rank (continued)

To see that they are orthogonal, compute

$$Av_i \cdot Av_j = v_i^T A^T A v_j = \lambda_j v_i^T v_j = 0$$

since the v_i are orthogonal. The Av_i also all belong to the column space of A.

Suppose that y is any vector in the column space of A. Then $y = Ax$ for some x, and

$$x = \sum_{i=1}^{n} (x \cdot v_i) v_i.$$

Apply A to this and since $Av_k = 0$ for $k > r$, we see that Ax is in the span of $Av_1, \dots, Av_r$.

So $Av_1, \dots, Av_r$ are orthogonal (hence linearly independent) and span the column space of A.

The SVD

Suppose that A is an $m \times n$ matrix of rank r. Then there exists an $m \times n$ matrix Σ which is "diagonal" in the sense that it looks like this:

$$\Sigma = \begin{array}{ll} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} & \begin{array}{l} \\ \leftarrow m - r \text{ rows} \end{array} \\ \qquad\quad \uparrow\ n - r \text{ columns} & \end{array}$$

Figure 1: "Diagonal" Matrix for SVD

where D is a truly diagonal $r \times r$ matrix whose entries are the nonzero singular values of A (in descending order), and orthogonal matrices U of size $m \times m$ and V of size $n \times n$ such that

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

Constructing the SVD

1. Let $u_i = \frac{Av_i}{\|Av_i\|} = \sigma_i^{-1} Av_i$ for $i = 1, \ldots, r$. This gives an orthonormal family. Extend this to an orthonormal basis $u_1 \ldots, u_m$ of $\mathbf{R}^m$.
2. Let U be the matrix whose columsn are the u_i and V be the matrix whose columns are the v_i.
3. Notice that AV has columns $\sigma_i u_i$ for $i = 1, \ldots, r$ and the rest zero. That's what you get if you compute $U\Sigma$.
4. So $AV = U\Sigma$ or $A = U\Sigma V^{-1} = U\Sigma V^T$.