Singular Values

Definition 14.1

Let A be a real $k \times n$ matrix. The singular values of A are the (positive) square roots of the eigenvalues of $A^T A$ and are denoted by $\sigma_1, \sigma_2, \ldots, \sigma_n$. It is conventional to arrange the singular values so that $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n$.

Singular values are, by construction, non-negative. For a symmetric matrix A, $A^T A = A^2$ and so

$$\mu_i = \lambda_i^2$$

where λ_i are the eigenvalues of A. Therefore the singular values of a symmetric matrix are

$$\sigma_i = \sqrt{\lambda_i^2} = |\lambda_i|$$

(note that the eigenvalues of A are real but may be negative).

Singular Values

Example

Find the singular value(s) of

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{w}^T = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Solution: We have

$$\mathbf{v}^T \mathbf{v} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \end{bmatrix}$$

and so the "eigenvalue" of $\mathbf{v}^T \mathbf{v}$ is 3. Therefore the singular value is

$$\sigma_1 = \sqrt{3}.$$

Singular Values

Now

$$\mathbf{w}^T \mathbf{w} = \mathbf{v}\,\mathbf{v}^T = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

The eigenvalues of this matrix (check) are 3, 0 and 0 and so the singular values are

$$\sigma_1 = \sqrt{3}, \quad \sigma_2 = \sigma_3 = 0.$$

Singular Values

Since $A^T A$ is symmetric, there is an *orthonormal* basis for $\mathbf{R}^n$ consisting of eigenvectors of $A^T A$. Let this basis be given by $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ where we order the eigenvalues of $A^T A$ so that $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$. We have observed earlier that

$$\mu_i = \|A\,\mathbf{v}_i\|^2$$

and so

$$\sigma_i = \sqrt{\mu_i} = \|A\,\mathbf{v}_i\|.$$

Now, for $i \neq j$,

$$(A\,\mathbf{v}_i) \cdot (A\,\mathbf{v}_j) = (A\,\mathbf{v}_i)^T\, A\,\mathbf{v}_j = \mathbf{v}_i^T A^T A \mathbf{v}_j = \mu_j\,(\mathbf{v}_i \cdot \mathbf{v}_j) = 0$$

since the eigenvectors $\mathbf{v}_i$ are orthogonal. Suppose the *non-zero* singular values of A are

$$\sigma_1 \geqslant \sigma_2 \cdots \geqslant \sigma_r > 0$$

with $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.

Singular Value Decomposition

Since $\sigma_i = \|A\,\mathbf{v}_i\|$,

$$\mathbf{u}_i = \frac{1}{\sigma_i} A\,\mathbf{v}_i$$

for $i = 1, 2, \ldots, r$ will be *unit* vectors which are mutually orthogonal; that is, they form an orthonormal set. Furthermore $\|A\,\mathbf{v}_i\| = \sigma_i = 0$; that is

$$A\,\mathbf{v}_i = \mathbf{0}$$

for $i = r+1, r+2, \ldots, n$. Therefore

$$A\begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_r & \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n\end{bmatrix} = \begin{bmatrix}\sigma_1\,\mathbf{u}_1 & \cdots & \sigma_r\,\mathbf{u}_r & \mathbf{0} & \cdots & \mathbf{0}\end{bmatrix}.$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbf{R}^n$,

$$V = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{bmatrix}$$

is an orthogonal $n \times n$ matrix. Note that $\mathbf{u}_i \in \mathbf{R}^k$. If $r < k$ we use Gram Schmidt to construct an orthonormal basis for $\mathbf{R}^k$ that includes $\mathbf{u}_i$; that is, we add $\mathbf{u}_j$, $j = r+1, r+2, \ldots, k$ such that

$$U = \begin{bmatrix}\mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r & \mathbf{u}_{r+1} & \ldots & \mathbf{u}_k\end{bmatrix}$$

is an orthogonal $k \times k$ matrix.

Singular Value Decomposition

We now have

$$A\,V = U\,\Sigma$$

where Σ is the $k \times n$ "diagonal" matrix

$$\Sigma = \begin{bmatrix}\sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0\\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0\end{bmatrix} = \begin{bmatrix}D & O\\ O & O\end{bmatrix}$$

with

$$D = \begin{bmatrix}\sigma_1 & 0 & \cdots & 0\\ 0 & \sigma_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sigma_r\end{bmatrix}.$$

Singular Value Decomposition

Note that Σ is the same size as A. Since V is orthogonal, we have

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

This is called the *singular value decomposition (SVD)* of A. The columns of U are called the *left singular vectors* and the columns of V are called the *right singular vectors*. Formally we have

Singular Value Decomposition

Theorem 14.2 (Singular Value Decomposition)

Let A be a real $k \times n$ matrix with singular values $\sigma_1 \geqslant \sigma_2 \cdots \geqslant \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$. Let Σ be the $k \times n$ matrix

$$\Sigma = \begin{bmatrix} D & O \\ O & O \end{bmatrix}$$

with

$$D = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}.$$

Then there exists an orthogonal $n \times n$ matrix V and an orthogonal $k \times k$ matrix U such that

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

The matrix Σ is fixed but the matrices U and V are *not* unique.

Singular Value Decomposition

For a symmetric matrix A, the SVD reduces to an orthogonal diagonalization. Recall that the singular values of A are

$$\sigma_i = |\lambda_i|$$

where λ_i are the eigenvalues of A. The eigenvectors $\mathbf{v}_i$ of $A^T A = A^2$ are the eigenvectors $\mathbf{q}_i$ of A and so

$$V = Q.$$

Furthermore

$$\mathbf{u}_i = \frac{1}{|\lambda_i|} A\, \mathbf{q}_i = \frac{1}{|\lambda_i|} \lambda_i\, \mathbf{q}_i = \mathrm{sign}(\lambda_i)\, \mathbf{q}_i.$$

Therefore

Singular Value Decomposition

$$A = U \Sigma V^T = \begin{bmatrix} \mathrm{sign}(\lambda_1)\, \mathbf{q}_1 & \cdots & \mathrm{sign}(\lambda_n)\, \mathbf{q}_n \end{bmatrix} \begin{bmatrix} |\lambda_1| & 0 & \cdots & 0 \\ 0 & |\lambda_2| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\lambda_n| \end{bmatrix} Q^T$$

$$= Q \begin{bmatrix} \mathrm{sign}(\lambda_1)\, |\lambda_1| & 0 & \cdots & 0 \\ 0 & \mathrm{sign}(\lambda_2)\, |\lambda_2| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{sign}(\lambda_n)\, |\lambda_n| \end{bmatrix} Q^T$$

$$= Q \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} Q^T$$

Singular Value Decomposition

Example

Find a SVD for

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

Solution: We have

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

and its eigenvalues are $\mu_1 = 2$, $\mu_2 = 1$ and $\mu_3 = 0$. Since the eigenvalues are distinct, the eigenvectors will be mutually orthogonal. The normalized eigenvectors are

$$\mathbf{v}_1 = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{\sqrt{2}}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Singular Value Decomposition

Therefore

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \end{bmatrix}.$$

The singular values of A are $\sigma_1 = \sqrt{2}$, $\sigma_2 = 1$ and $\sigma_3 = 0$. Thus Σ is the 2×3 matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Finally

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A \mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A \mathbf{v}_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Singular Value Decomposition

These vectors already form an orthonormal basis for $\mathbf{R}^2$ and so

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This yields the SVD

$$A = U \Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

as expected.

Singular Value Decomposition

Example

Find a SVD for

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

Solution: We have

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

and its eigenvalues are $\mu_1 = 3$ and $\mu_2 = 1$. Since the eigenvalues are distinct, the eigenvectors will be mutually orthogonal. The normalized eigenvectors are

$$\mathbf{v}_1 = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{\sqrt{2}}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Singular Value Decomposition

Therefore

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

The singular values of A are $\sigma_1 = \sqrt{3}$ and $\sigma_2 = 1$. Thus Σ is the 3×2 matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Now

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{\sqrt{6}}{6} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A \mathbf{v}_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{\sqrt{2}}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{\sqrt{2}}{2} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

Singular Value Decomposition

In this case $\mathbf{u}_1$ and $\mathbf{u}_2$ do not form a basis for $\mathbf{R}^3$. We choose a linearly independent vector and then use Gram Schmidt to construct an orthonormal basis for $\mathbf{R}^3$. Since $\mathbf{u}_1$ and $\mathbf{u}_2$ are already orthogonal, we only need the last step of the Gram Schmidt process. Clearly

$$\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

is linearly independent to $\mathbf{u}_1$ and $\mathbf{u}_2$ and so

$$\mathbf{e}_3 - (\mathbf{e}_3 \cdot \mathbf{u}_1)\,\mathbf{u}_1 - (\mathbf{e}_3 \cdot \mathbf{u}_2)\,\mathbf{u}_2 = \mathbf{e}_3 - \frac{\sqrt{6}}{6}\,\mathbf{u}_1 - \frac{\sqrt{2}}{2}\,\mathbf{u}_2 = \frac{1}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

will be orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$. Normalizing this vector we have

Singular Value Decomposition

$$\mathbf{u}_3 = \frac{\sqrt{3}}{3} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

and so

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix}.$$

This yields the SVD

$$A = U\Sigma V^T = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

Singular Value Decomposition

There is a form of the SVD which is analogous to the spectral decomposition of a symmetric matrix. We have

$$A = U\Sigma V^T = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_k \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ & O & & O \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_k \end{bmatrix} \begin{bmatrix} \sigma_1 \mathbf{v}_1^T \\ \vdots \\ \sigma_r \mathbf{v}_r^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T.$$

Singular Value Decomposition

Theorem 14.3 (Outer Product form of the SVD)

Let A be a real $k \times n$ matrix with singular values $\sigma_1 \geqslant \sigma_2 \cdots \geqslant \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$. Let $\mathbf{u}_1, \ldots, \mathbf{u}_r$ be the left singular vectors and $\mathbf{v}_1, \ldots, \mathbf{v}_r$ be the right singular vectors corresponding to the non-zero singular values. Then

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T.$$

Note that, in this form, there is no need to compute the singular vectors corresponding to the zero singular values. However the singular vectors span the fundamental subspaces of A.

Singular Value Decomposition

Theorem 14.4

Let $A = U \Sigma V^T$ be a singular value decomposition of a $k \times n$ matrix A. Let $\sigma_1, \ldots, \sigma_r$ be the non-zero singular values of A. Then

(a) *The rank of A is r.*

(b) $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ *is an orthonormal basis for* $\operatorname{col}(A)$.

(c) $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_k\}$ *is an orthonormal basis for* $\operatorname{null}\left(A^T\right)$.

(d) $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ *is an orthonormal basis for* $\operatorname{row}(A)$.

(e) $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ *is an orthonormal basis for* $\operatorname{null}(A)$.

Proof: See Poole Theorem 7.15.

In particular, this theorem implies that a $n \times n$ matrix is invertible if and only if it has no zero singular values; that is $r = n$.

Compression

The other product form of an SVD implies that we only need to store r vectors $\mathbf{u}_i \in \mathbf{R}^k$, r vectors $\mathbf{v}_i \in \mathbf{R}^n$ and the r non-zero singular values σ_i in order to reconstruct the matrix A. In total, there are

$$rk + rn + r = r\,(k + n + 1)$$

numbers that we need to store compare to kn numbers that would be needed if we stored A directly. In fact we can do slightly better. If we store $\sigma_i\,\mathbf{u}_i = A\,\mathbf{v}_i$ then the storage is reduced to

$$r\,(k + n)\,.$$

If r is small (that is, the rank is small) then

$$r\,(k + n) << kn$$

and we have achieved *loss-less compression*. In particular, for a square $n \times n$ matrix, we require

$$r < \tfrac{1}{2}\,n$$

in order to achieve loss-less compression.

Compression

In the case of loss-less compression, we can reconstruct A *exactly*. However if r is large, we might want to *approximate* A by a matrix of *smaller rank*. This will then achieve *lossy compression* since we will not be able to reconstruct A exactly.

Each term

$$\sigma_i\,\mathbf{u}_i\,\mathbf{v}_i^T$$

in the SVD is a rank 1 matrix. Since we have ordered the singular values in descending size, we call

$$A_1 = \sigma_1\,\mathbf{u}_1\,\mathbf{v}_1^T$$

the *rank 1 approximation to A*. Similarly, for $s < r$, we call

$$A_s = \sigma_1\,\mathbf{u}_1\,\mathbf{v}_1^T + \cdots + \sigma_s\,\mathbf{u}_s\,\mathbf{v}_s^T$$

the *rank s* approximation to A.

Compression

In typical applications (say image compression), there is a fair range of singular values. By ignoring the small non-zero singular values, we achieve a lower rank approximation to A that preserves *most* of the information contained in A (remember that the terms that we are ignoring will be small compared to the terms that we are keeping). The art (or engineering!) is to choose s as small as possible (to maximize the compression) but not too small so that we loose essential detail in A.

Compression

MATLAB

MATLAB has the function `svd` to compute the singular decomposition of a matrix A.

- `s = svd(A)` (or, simply, `svd(A)`) will return a *vector* of the singular values of A (in order of descending size).
- `[U,S,V] = svd(A)` will return the three factors of a SVD for A.

Consider the matrices

$$A = \begin{bmatrix} 8.1650 & -0.0041 & -0.0041 \\ 4.0825 & -3.9959 & 4.0042 \\ 4.0825 & 4.0042 & -3.9959 \end{bmatrix}, \qquad B = \begin{bmatrix} 8.17 & 0 & 0 \\ 4.08 & -4.00 & 4.00 \\ 4.08 & 4.00 & -4.00 \end{bmatrix}$$

where B has been obtained from A by rounding to two decimal places. MATLAB gives the singular values

```
>> [svd(A),svd(B)]
ans =
   10.0000   10.0000
    8.0000    8.0000
    0.0100    0.0000
```

SINGULAR VALUE DECOMPOSITION

We see that A has rank 3 and so is invertible but B has rank 2 and thus is not invertible. In applications, it is often assumed that small singular values are the result of round-off error and that the actual value should be zero. In this way "noise" can be filtered out. A SVD for A is

```
>> [U,S,V]=svd(A)
U =
   -0.8165   -0.0000   -0.5774
   -0.4082   -0.7071    0.5774
   -0.4082    0.7071    0.5774
S =
   10.0000         0         0
         0    8.0000         0
         0         0    0.0100
V =
   -1.0000         0         0
         0    0.7071    0.7071
         0   -0.7071    0.7071
```

Singular Value Decomposition

If we assume that the singular value 0.01 is the result of round-off error then

```
>> U*diag([10,8,0])*V'
ans =
    8.1650   -0.0000    0.0000
    4.0825   -4.0000    4.0000
    4.0825    4.0000   -4.0000
```

that is,

$$U \begin{bmatrix} 10 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T = B.$$

Singular Value Decomposition

Theorem 14.5

Let A be an invertible matrix with a singular value decomposition $A = U\Sigma V^T$. Then A^{-1} has a singular value decomposition

$$A^{-1} = V\Sigma^{-1}U^T.$$

Proof: Exercise.

Note that this theorem implies that the singular values of A^{-1} are

$$\frac{1}{\sigma_n} \geqslant \frac{1}{\sigma_{n-1}} \geqslant \cdots \geqslant \frac{1}{\sigma_1}.$$

Condition Number

Recall that

$$\|A\|_2 = \max_{\|\mathbf{x}\|=1} \|A\,\mathbf{x}\|.$$

Let $A = U\,\Sigma\,V^T$ be a SVD for A. Since U is orthogonal, we have $\|U\,\mathbf{y}\| = \|\mathbf{y}\|$ for any $\mathbf{y} \in \mathbf{R}^k$ and so

$$\|A\,\mathbf{x}\| = \|U\,\Sigma\,V^T\,\mathbf{x}\| = \|\Sigma\,V^T\,\mathbf{x}\|$$

(set $\mathbf{y} = \Sigma\,V^T\,\mathbf{x}$). Since V (and therefore V^T) is orthogonal, for $\|\mathbf{x}\| = 1$ we have

$$\|V^T\,\mathbf{x}\| = \|\mathbf{x}\| = 1.$$

Therefore, setting $\mathbf{y} = V^T\,\mathbf{x}$ we have

$$\|A\|_2 = \max_{\|\mathbf{x}\|=1} \|A\,\mathbf{x}\| = \max_{\|\mathbf{y}\|=1} \|\Sigma\,\mathbf{y}\| = \sigma_1.$$

From the above theorem we also note that, if A is invertible,

$$\|A^{-1}\|_2 = \frac{1}{\sigma_n}.$$

Condition Number

Therefore the condition number (using the 2-norm) for a matrix is

$$\text{cond}_2\,(A) \equiv \|A^{-1}\|_2\,\|A\|_2 = \frac{\sigma_1}{\sigma_n}.$$

For the matrix A in the above example, we have

$$\text{cond}_2\,(A) = \frac{10}{0.01} = 1000.$$

Pseudoinverse

We have defined the pseudoinverse for a matrix A that has linearly independent columns. We can now extend this definition to *any* matrix; in particular, to non-invertible square matrices.

Definition 14.6

Let $A = U\Sigma V^T$ be a SVD of a $k \times n$ matrix A with

$$\Sigma = \begin{bmatrix} D & O \\ O & O \end{bmatrix}$$

and D is the $r \times r$ diagonal matrix with the non-zero singular values $\sigma_1 \geqslant \sigma_2 \cdots \geqslant \sigma_r > 0$of A on its diagonal. The pseudoinverse or Moore-Penrose inverse of A is the $n \times k$ matrix

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

where Σ^+ is the $n \times k$ matrix

$$\Sigma^+ = \begin{bmatrix} D^{-1} & O^T \\ O^T & O^T \end{bmatrix}.$$

Pseudoinverse

This new definition reduces to the old one in the case of a $k \times n$ matrix A with linearly independent columns. Suppose $A = U\Sigma V$ is a SVD for A. Then

$$A^T A = V\Sigma^T U^T U\Sigma V^T = V\left(\Sigma^T \Sigma\right)V^T$$

since U is orthogonal. Since A has linearly independent columns, it has no non-zero singular values $(r = n)$ and so Σ is the $k \times n$ matrix

$$\Sigma = \begin{bmatrix} D \\ O \end{bmatrix}$$

and so

$$\Sigma^T \Sigma = \begin{bmatrix} D \\ O^T \end{bmatrix} \begin{bmatrix} D & O \end{bmatrix} = D^2.$$

Therefore

$$\left(A^T A\right)^{-1} = \left(V D^2 V^T\right)^{-1} = V^{-T} D^{-2} V^{-1}.$$

PSEUDOINVERSE

Since V is orthogonal,

$$\begin{aligned} \left(A^T A\right)^{-1} A^T &= V D^{-2} V^T V \Sigma^T U^T \\ &= V D^{-2} \begin{bmatrix} D & O^T \end{bmatrix} U^T \\ &= V \begin{bmatrix} D^{-1} & O^T \end{bmatrix} V^T \\ &= V \Sigma^+ U^T. \end{aligned}$$

PSEUDOINVERSE

Example

Find the pseudoinverse of

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

Solution: Note that A does not have linearly independent columns. From the example above, we have the SVD

$$A = U \Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for A. Therefore

$$\Sigma^+ = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Pseudoinverse

and so

$$A^+ = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}.$$

Pseudoinverse

Example

Find the pseudoinverse of

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

Solution: From above, we have the SVD

$$A = U\Sigma V^T = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

Thus

$$\Sigma^+ = \begin{bmatrix} \frac{\sqrt{3}}{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Pseudoinverse

and so
$$A^+ = V\Sigma^+ U^T = \frac{1}{3}\begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Least Squares

We have seen that there is a unique least squares solution to $A\mathbf{x} = \mathbf{b}$ if and only if A has linearly independent columns. In the case where A does not have linearly independent columns, there are infinitely many solutions. In this case, we will ask for the solution of *minimum length*; that is, closest to the origin.

Theorem 14.7

The system $A\mathbf{x} = \mathbf{b}$ has a unique least squares solution of minimal length. It is given by
$$\bar{\mathbf{x}} = A^+\mathbf{b}.$$

Proof: See Poole Theorem 7.18.

Least Squares

Example

Find the minimal least squares solution of

$$x_1 + x_2 = 0$$
$$x_1 + x_2 = 1.$$

Solution: The coefficient matrix is

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

which has a SVD (check)

$$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}$$

(the fact that $U = V^T$ should not be surprising; A is symmetric).

Least Squares

Thus

$$A^+ = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

and so the least squares solution of minimal length is

$$\bar{\mathbf{x}} = A^+ \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$