Lecture 11

- The SVD: extending diagonalisation to non-square matrices

The Singular Value Decomposition

Notes:

- Diagonalization (ie the eigenvalue problem) plays important role in applications **BUT** not all matrices can be factored in the form $\mathbf{A} = \mathbf{PDP^{-1}}$ with $\mathbf{D}$ diagonal.
- **However** a factorization (the **singular value decomposition**) $\mathbf{A} = \mathbf{QDP^{-1}}$ (with $\mathbf{D}$ diagonal) **is** possible for any $m \times n$ matrix $\mathbf{A}$.
- The $|\lambda|$ of a symmetric matrix $\mathbf{A}$ measures the amounts that $\mathbf{A}$ stretches or shrinks eigenvectors: if $\mathbf{Ax} = \lambda\mathbf{x}$ and $\|\mathbf{x}\| = 1$, then
$$\|\mathbf{Ax}\| = \|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\| = |\lambda|.$$
If the eigenvalue λ_1 is the eigenvalue with the greatest magnitude then a corresponding unit eigenvector $\mathbf{v}_1$ identifies the direction of the greatest stretching.
- The length of $A\mathbf{x}$ is maximized when $\mathbf{x} = \mathbf{v}_1$
- This description has an analog for rectangular matrices.
- This will lead to the singular value decomposition.

The Singular Value Decomposition

Example: If

$$\mathbf{A} = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$$

then the transformation $\mathbf{x} \to \mathbf{Ax}$ maps the unit sphere $\{\mathbf{x} : \|\mathbf{x}\| = 1\}$ in $\mathbb{R}^3$ onto an ellipse in $\mathbb{R}^2$.

Find a unit vector $\mathbf{x}$ at which the length $\|A\mathbf{x}\|$ is maximized and compute its length.

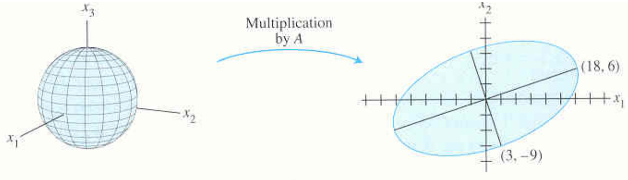


FIGURE 1 A transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$.

The Singular Value Decomposition

Solution: The quantity $\|A\mathbf{x}\|^2$ is maximized at the same $\mathbf{x}$ that maximizes $\|\mathbf{Ax}\|$. Also,

$$\begin{aligned}\|\mathbf{Ax}\|^2 &= (\mathbf{Ax})^T(\mathbf{Ax}) = \mathbf{x}^T\mathbf{A}^{\mathbf{T}}\mathbf{Ax} \\ &= \mathbf{x}^T(\mathbf{A}^{\mathbf{T}}\mathbf{A})\mathbf{x}\end{aligned}$$

Note that the matrix $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ is symmetric, since

$$(\mathbf{A}^{\mathbf{T}}\mathbf{A})^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}}\mathbf{A}^{\mathbf{TT}} = \mathbf{A}^{\mathbf{T}}\mathbf{A}.$$

So the problem is now familiar: maximize the quadratic form $\mathbf{x}^T(\mathbf{A}^{\mathbf{T}}\mathbf{A})\mathbf{x}$ subject to the constraint $\|\mathbf{x}\| = 1$.

By Theorem the maximum value is the greatest eigenvalue λ_1 of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ and the maximum value is attained at the unit eigenvector of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ corresponding to λ_1.

The Singular Value Decomposition

We have:

$$\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{pmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{pmatrix} \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}.$$

The eigenvalues of this matrix are $\lambda_1 = 360$, $\lambda_2 = 90$, $\lambda_3 = 0$. The corresponding unit eigenvectors are:

$$\mathbf{v}_1 = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ 1/3 \end{pmatrix}.$$

Maximum value (for $\mathbf{x}$ with $\|\mathbf{x}\| = 1$) of $\|A\mathbf{x}\|^2$ is 360, attained at $\mathbf{x} = \mathbf{v}_1$. Vector $A\mathbf{v}_1$ is a point on the ellipse furthest from $\mathbf{0}$:

$$A\mathbf{v}_1 = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 18 \\ 6 \end{pmatrix}.$$

The maximum value of $\|A\mathbf{x}\| = \sqrt{360} = 6\sqrt{10}$.

The Singular Value of Matrix **A**

- Let **A** be an $m \times n$ matrix. Then $\mathbf{A^T A}$ is symmetric and can be orthogonally diagonalized.
- Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$ consisting of unit eigenvectors of $\mathbf{A^T A}$, and $\lambda_1, \ldots, \lambda_n$ be the associated eigenvalues of $\mathbf{A^T A}$. Then

$$\begin{aligned}\|\mathbf{Av_i}\|^2 &= (\mathbf{Av_i})^T A\mathbf{v}_i = \mathbf{v}_i\mathbf{A^T A v_i} = \mathbf{v_i^T}(\lambda_\mathbf{i})\mathbf{v_i}\\ &= \lambda_i,\end{aligned}$$

  so the eigenvalues of $\mathbf{A^T A}$ are all nonnegative.
- We can always arrange them in descending order so that

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0.$$

The Singular Value of Matrix **A**

Definition: The **singular values** of **A** are the square roots of the eigenvalues $\lambda_1, \dots, \lambda_n$ of $\mathbf{A^T A}$, denoted by $\sigma_1, \dots, \sigma_n$ arranged in the descending order. So

$$\sigma_i = \sqrt{\lambda_i}, \ \textit{for} \ \ 1 \le i \le n.$$

Note: With the notation above, the singular values of A are the lengths of the vectors $\mathbf{Av_1}, \dots, \mathbf{Av_n}$.

The Singular Values of Matrix $\mathbf{A}$

Example: Let $\mathbf{A}$ be as in the last example: $\mathbf{A} = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$.

The eigenvalues of $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ are $\lambda_1 = 360$, $\lambda_2 = 90$, $\lambda_3 = 0$ therefore the singular values of A are

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \ \ \sigma_2 = \sqrt{90} = 3\sqrt{10}, \ \ \sigma_3 = 0.$$

The first singular value of $\mathbf{A}$ is the maximum of $\|\mathbf{Ax}\|$ subject to $\|\mathbf{x}\| = 1$, attained at $\mathbf{x} = \mathbf{v}_1$.
The second singular value of $\mathbf{A}$ is the maximum of $\|\mathbf{Ax}\|$ over all unit vectors orthogonal to $\mathbf{v}_1$ and this is attained at $\mathbf{x} = \mathbf{v}_2$.

The Singular Values of Matrix $\mathbf{A}$

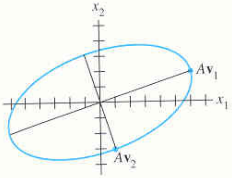


FIGURE 2

$$\mathbf{Av_2} = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 3 \\ -9 \end{pmatrix}.$$

This point is on the minor axis of the ellipse.

The Singular Value of Matrix $\mathbf{A}$

Theorem: Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A^T A}$, arranged so that the corresponding eigenvalues of $\mathbf{A^T A}$ satisfy $\lambda_1 \geq \ldots \geq \lambda_n$, and suppose $\mathbf{A}$ has r nonzero singular values. Then $\{\mathbf{Av_1}, \ldots, \mathbf{Av_r}\}$ is an orthogonal basis for $Col\, \mathbf{A}$ and rank $A = r$.

Proof: We have:

$$(\mathbf{Av_i})^T(\mathbf{Av_j}) = \mathbf{v}_i^T \mathbf{A^T A v_j} = \mathbf{v_i^T}(\lambda_\mathbf{j} \mathbf{v_j}) = \lambda_\mathbf{j}(\mathbf{v_i^T v_j})$$

$$= \begin{cases} 0, & i \neq j, \\ \lambda_j, & i = j. \end{cases}$$

Therefore $\{\mathbf{Av_1}, \ldots, \mathbf{Av_r}\}$ is an orthogonal set.

The lengths of the vectors $\{\mathbf{Av_1}, \ldots, \mathbf{Av_n}\}$ are the singular values of $\mathbf{A}$, of which the first r are strictly positive, and hence $\{\mathbf{Av_1}, \ldots, \mathbf{Av_r}\}$ are non-zero vectors.

Thus $\{\mathbf{Av_1}, \ldots, \mathbf{Av_r}\}$ is a linearly independent set of orthogonal vectors in the column space $Col\, \mathbf{A}$. We must show it spans $Col\, \mathbf{A}$.

Proof (cont'd):

For any $\mathbf{y}$ in $Col\,\mathbf{A}$ we have $\mathbf{y} = \mathbf{A}\mathbf{x}$, where

$$\mathbf{x} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$$

and so

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{x} \\ &= c_1\mathbf{A}\mathbf{v_1} + \ldots + c_r\mathbf{A}\mathbf{v_r} + c_{r+1}\mathbf{A}\mathbf{v}_{r+1} + \ldots + c_n\mathbf{A}\mathbf{v_n} \\ &= c_1\mathbf{A}\mathbf{v_1} + \ldots + c_r\mathbf{A}\mathbf{v_r} + c_{r+1}\mathbf{0} + \ldots + c_n\mathbf{0}. \end{aligned}$$

Therefore $\mathbf{y}$ is in $Span\{\mathbf{A}\mathbf{v_1}, \ldots, \mathbf{A}\mathbf{v_r}\}$ which means that the set $\{\mathbf{A}\mathbf{v_1}, \ldots, \mathbf{A}\mathbf{v_r}\}$ is an orthogonal basis for $Col\,\mathbf{A}$.

Hence $rank\,\mathbf{A} = dim\,Col\,\mathbf{A} = r$.

The Singular Value Decomposition

The decomposition of $\mathbf{A}$ involves an $m \times n$ "diagonal" matrix $\mathbf{\Sigma}$ of the form

$$\mathbf{\Sigma} = \begin{pmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

where $\mathbf{D}$ is an $r \times r$ diagonal matrix for some r not exceeding the smaller of m and n. The second line in $\mathbf{\Sigma}$ contains $m - r$ rows. The second column in $\mathbf{\Sigma}$ contains $n - r$ columns.

Theorem: <u>Singular Value Decomposition</u>
Let $\mathbf{A}$ be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix $\mathbf{\Sigma}$ for which the diagonal entries in $\mathbf{D}$ are the first r singular values of $\mathbf{A}$ $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$, and there exist an $m \times m$ orthogonal matrix $\mathbf{U}$ and an $n \times n$ orthogonal matrix $\mathbf{V}$ such that

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

This factorisation is called the **singular value factorisation** of $\mathbf{A}$.

The Singular Value Decomposition

Notes:

- The matrices $\mathbf{U}$ and $\mathbf{V}$ in $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathbf{T}}$ are not uniquely determined by $\mathbf{A}$ but the diagonal entries in $\boldsymbol{\Sigma}$ are uniquely determined by the singular values of $\mathbf{A}$.
- The columns of $\mathbf{U}$ are called **left singular vectors** of $\mathbf{A}$ and the columns of $\mathbf{V}$ are called the **right singular vectors** of $\mathbf{A}$.

The Singular Value Decomposition

Proof: Let λ_i be the eigenvalues and $\mathbf{v}_i$ be corresponding eigenvectors of $\mathbf{A^T A}$. Then $\{\mathbf{Av_1}, \ldots, \mathbf{Av_r}\}$ is an orthogonal basis for *Col* $\mathbf{A}$.

We normalize each $\mathbf{Av_i}$ to obtain an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ for *Col* $\mathbf{A}$:

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

from which we obtain

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i, \;\; 1 \leq i \leq r.$$

The set $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ can be extended, if necessary, to an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ for $\mathbb{R}^m$ by adding suitable vectors from the orthogonal complement of $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$.

The Singular Value Decomposition

Now let

$$\mathbf{U} = (\mathbf{u}_1 \ \dots \ \mathbf{u}_m), \quad \text{and} \quad \mathbf{V} = (\mathbf{v}_1 \ \dots \ \mathbf{v}_n).$$

By construction, $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices. Now

$$\mathbf{AV} = (A\mathbf{v}_1 \ \dots \ A\mathbf{v}_r \ \ \mathbf{0} \dots \mathbf{0}) = (\sigma_1 \mathbf{u}_1 \dots \sigma_r \mathbf{u}_r \ \ \mathbf{0} \dots \ \mathbf{0})$$

and

$$\mathbf{U\Sigma} = (\sigma_1 \mathbf{u}_1 \ \dots \ \sigma_r \mathbf{u}_r \ \ \mathbf{0} \dots \mathbf{0}) = \mathbf{AV},$$

that is,

$$\mathbf{AV} = \mathbf{U\Sigma}.$$

But $\mathbf{V}$ is an orthogonal matrix, so $\mathbf{V^T V} = \mathbf{I}$ and

$$\mathbf{AVV^T} = \mathbf{U\Sigma V^T},$$

and so

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

Example: Construct a singular value decomposition of $\mathbf{A}$:

$$\mathbf{A} = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}.$$

Solution: Step 1: Construct an orthogonal diagonalization of $\mathbf{A}^\mathbf{T}\mathbf{A}$
We need the eigenvalues and corresponding eigenvectors of $\mathbf{A}^\mathbf{T}\mathbf{A}$.
We have already calculated them in previous examples. They are (in descending order): 360, 90, 0.
The corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}, \ \mathbf{v}_2 = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \ \mathbf{v}_3 = \begin{pmatrix} 2/3 \\ -2/3 \\ 1/3 \end{pmatrix}.$$

Thus

$$\mathbf{V} = (\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \mathbf{v}_3) = \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix}.$$

The Singular Value Decomposition

Step 2: The singular values of $\mathbf{A}$ are:

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}, \quad \sigma_3 = 0.$$

The nonzero values are diagonal values of $\mathbf{D}$:

$$\mathbf{D} = \begin{pmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{pmatrix},$$

$$\boldsymbol{\Sigma} = (\mathbf{D} \;\; \mathbf{0}) = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}.$$

The Singular Value Decomposition

Step 3: Construct $\mathbf{U}$. When $\mathbf{A}$ has rank r the first r columns of $\mathbf{U}$ are normalized vectors obtained from $\mathbf{Av_1}, \ldots, \mathbf{Av_r}$. $\mathbf{A}$ has two nonzero singular values so *rank* $\mathbf{A} = 2$ and

$$\|\mathbf{Av_1}\| = \sigma_1, \quad \|\mathbf{Av_2}\| = \sigma_2.$$

Thus the columns of $\mathbf{U}$ are

$$\mathbf{u}_1 = \frac{1}{\sigma_1}\mathbf{Av_1} = \frac{1}{6\sqrt{10}}\begin{pmatrix} 18 \\ 6 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix}$$
$$\mathbf{u}_2 = \frac{1}{\sigma_2}\mathbf{Av_2} = \frac{1}{3\sqrt{10}}\begin{pmatrix} 3 \\ -9 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{pmatrix}.$$

The set $\{\mathbf{u}_1, \mathbf{u}_2\}$ is already a basis for $\mathbb{R}^2$. No additional vectors are needed for $\mathbf{U}$ and so $\mathbf{U} = (\mathbf{u_1} \;\; \mathbf{u_2})$.

The Singular Value Decomposition

Thus the singular value decomposition is

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

$$= \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{-2}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{pmatrix}$$

The Singular Value Decomposition

Example: Find a singular value decomposition for $\mathbf{A}$ given by

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}.$$

Solution: $\mathbf{A} = \mathbf{U\Sigma V^T}$. First we calculate $\mathbf{A^T A}$

$$\mathbf{A^T A} = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$$

Eigenvalues: $\lambda_1 = 18$ and $\lambda_2 = 0$ with corresponding eigenvectors:

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

These unit vectors form columns of $\mathbf{V}$:

$$\mathbf{V} = (\mathbf{v}_1 \ \mathbf{v}_2) = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

The Singular Value Decomposition

Singular values of $\mathbf{A}$:

$$\sigma_1 = \sqrt{18} = 3\sqrt{2}, \quad \sigma_2 = 0.$$

Since there is only one nonzero singular value the matrix $\mathbf{D}$ is of order 1×1:

$$D = (3\sqrt{2}).$$

The matrix $\mathbf{\Sigma}$ is the same size as $\mathbf{A}$:

$$\mathbf{\Sigma} = \begin{pmatrix} D & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

To construct $\mathbf{U}$ we first calculate $\mathbf{Av_1}$ and $A\mathbf{v}_2$:

$$\mathbf{Av_1} = \begin{pmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{pmatrix}, \ \mathbf{Av_2} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The Singular Value Decomposition

The only column found for $\mathbf{U}$ so far is therefore

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{3\sqrt{2}} = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}.$$

The other columns must be found by extending $\{\mathbf{u}_1\}$ to an orthogonal basis for $\mathbb{R}^3$.
We need two orthonormal vectors orthogonal to $\mathbf{u}_1$.
Each vector must satisfy $\mathbf{u}_1^T\mathbf{x} = 0$. This is equivalent to
$x_1 - 2x_2 + 2x_3 = 0$.
Solution is
$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{w}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$

The Singular Value Decomposition

We apply the Gram-Schmidt process to $\{\mathbf{w_1}, \mathbf{w_2}\}$ to get

$$\mathbf{u}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} -2/\sqrt{45} \\ 0/\sqrt{45} \\ 5/\sqrt{45} \end{pmatrix}.$$

Thus $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3]$ and the singular value decomposition is

$$\mathbf{A} = \begin{pmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

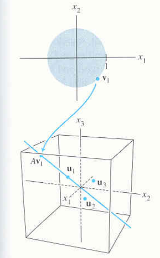


FIGURE 3

Bases for Fundamental Subspaces

Let

> $\mathbf{A}$ *be an* $m \times n$ *matrix;*
> $\mathbf{u}_1, \ldots, \mathbf{u}_m$ *be the left singular vectors;*
> $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *be the right singular vectors;*
> $\sigma_1, \ldots, \sigma_r$ *be the singular values;*
> r *be the rank of* $\mathbf{A}$*;*

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for *Col* $\mathbf{A}$.

Also recall that $(\mathit{Col}\,\mathbf{A})^{\perp} = \mathit{Nul}\,\left(\mathbf{A}^{\mathbf{T}}\right)$, and so $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$ is an orthonormal basis for $\mathit{Nul}\,\left(\mathbf{A}^{T}\right)$.

Since $\|\mathbf{A}\mathbf{v}_\mathbf{i}\| = \mathbf{0}$ for $i > r$, the size of *Nul* $\mathbf{A}$ is $n - r$ and so the set $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for *Nul* $\mathbf{A}$.

Finally, since $(\mathit{Nul}\,\mathbf{A})^{\perp} = \mathit{Col}\,\left(\mathbf{A}^{\mathbf{T}}\right) = \mathit{Row}\,\mathbf{A}$ so $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthonormal basis for *Row* $\mathbf{A}$.

Bases for Fundamental Subspaces

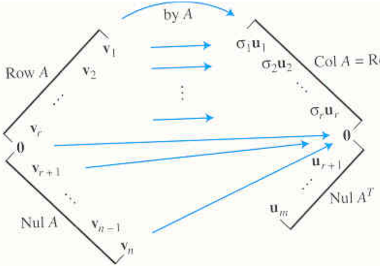
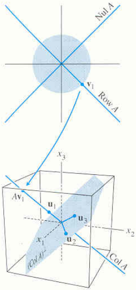


FIGURE 4 The four fundamental subspaces and the action of A.

ur fundamental subspaces and the concept of singular values pro



The fundamental subspaces in Example 4.

$$\mathbf{u}_1 = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1\sqrt{5}/ \\ 0 \end{pmatrix}, \quad \mathbf{u}_3 = \begin{pmatrix} -2/\sqrt{45} \\ 0/\sqrt{45} \\ 5/\sqrt{45} \end{pmatrix}$$

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

The Singular Value Decomposition

The SVD of $\mathbf{A}$ is $\mathbf{A} = \mathbf{U\Sigma V^T}$. We can write this as

$$\mathbf{A} = (\mathbf{u}_1 \ \ldots \ \mathbf{u}_m) \begin{pmatrix} \sigma_1 & \ldots & 0 & 0 & \\ 0 & \ddots & 0 & 0 & \ldots \\ 0 & 0 & \sigma_r & 0 & \\ 0 & 0 & 0 & 0 & \\ & & \vdots & & \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{pmatrix}$$

$$= \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T.$$

Notes:

- Original matrix required $m \times n$ floating point numbers to be saved — this expansion only $m \times r + n \times r + r = r(m + n + 1)$.
- Usually some of the singular values are very small therefore

  $$\mathbf{A}_k \approx \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T,$$

  where $k < r$. (k is the **rank** of the approximation.)
- This is why SVD-based image compression / dimension reduction works!

The Singular Value Decomposition

Rank 4 Rank 10 Rank 20 Rank 50 Rank 128

Figure 8.6.3

k is the rank of the approximation.