Fundamentals of linear algebra

- Approximate solutions: least-squares
- Revision: eigenvalues and eigenvectors
- Diagonalisation and orthogonal diagonalisation of matrices
- Spectral decomposition
- Quadratic forms
- Singular value decomposition

Wednesday, 29 May 2019

QR factorisation of matrices

Theorem: *QR factorisation*

An $m \times n$ matrix $\mathbf{A}$ with linearly independent columns can be factorised as $\mathbf{A} = \mathbf{QR}$, where $\mathbf{Q}$ is an $m \times n$ matrix with columns forming an orthonormal basis for $\operatorname{Col} \mathbf{A}$, and $\mathbf{R}$ is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

Note: Since $\mathbf{Q}$ is an orthonormal matrix, $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$.

Thus $\mathbf{Q}^\top \mathbf{A} = \mathbf{Q}^\top (\mathbf{QR}) = (\mathbf{Q}^\top \mathbf{Q})\mathbf{R} = \mathbf{IR} = \mathbf{R}$, so $\mathbf{R} = \mathbf{Q}^\top \mathbf{A}$.

Proof: by construction; see the previous lecture

QR factorisation of matrices

Example:

Find a QR decomposition of: $\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Solution: Earlier we have found an orthogonal basis for $\operatorname{Col}\mathbf{A}$ as

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}.$$

Upon normalisation we obtain

$$\mathbf{Q} = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}.$$

QR factorisation of matrices

$$\mathbf{R} = \mathbf{Q}^\top \mathbf{A} = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1//\sqrt{6} \end{bmatrix}^\top \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{12} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}.$$

So the QR decomposition is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$$

Least-squares solutions

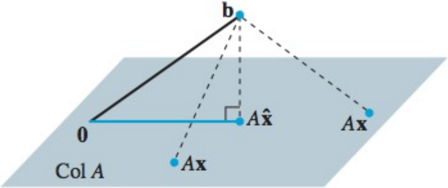
Definition: For an $m \times n$ matrix $\mathbf{A}$ and given $\mathbf{b} \in \mathbb{R}^m$, a *least-squares solution* of $\mathbf{A}\mathbf{x} = \mathbf{b}$ is $\hat{\mathbf{x}} \in \mathbb{R}^n$ such that $\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\| \leqslant \|\mathbf{b} - \mathbf{A}\mathbf{x}\| \quad \forall \mathbf{x} \in \mathbb{R}^n$.

Take $\hat{\mathbf{b}} = \mathrm{proj}_{(\mathrm{Col}\,\mathbf{A})}\,\mathbf{b}$, then $\hat{\mathbf{b}} \in \mathrm{Col}\,\mathbf{A}$ and $\exists\, \hat{\mathbf{x}} :\ \mathbf{A}\hat{\mathbf{x}} = \hat{\mathbf{b}}$.

In $\mathrm{Col}\,\mathbf{A}$, $\hat{\mathbf{b}}$ is the closest to $\mathbf{b}$, so $\hat{\mathbf{x}}$ is a least-square solution.

Via orthogonal decomposition, $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to $\mathrm{Col}\,\mathbf{A}$, so

$$
\begin{aligned}
\mathbf{A}^\top(\mathbf{b} - \hat{\mathbf{b}}) &= \mathbf{0} \\
\mathbf{A}^\top(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) &= \mathbf{0} \\
\mathbf{A}^\top\mathbf{b} - \mathbf{A}^\top\mathbf{A}\hat{\mathbf{x}} &= \mathbf{0} \\
\mathbf{A}^\top\mathbf{A}\hat{\mathbf{x}} &= \mathbf{A}^\top\mathbf{b}
\end{aligned}
$$



System $\mathbf{A}^\top\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top\mathbf{b}$ is the *normal system* for $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Its non-empty solution (set) is the least-squares solution (set).

Least-squares solutions

Example: Find a least-squares solution of the inconsistent system

$$\mathbf{A} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution:

$$\mathbf{A}^\top \mathbf{A} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$\mathbf{A}^\top \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \hat{\mathbf{x}} = \begin{bmatrix} 19 \\ 11 \end{bmatrix} \qquad \Rightarrow \qquad \hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

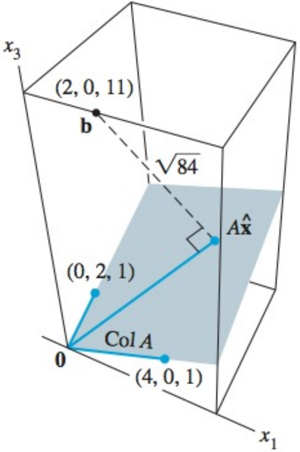
Least-squares solutions

Having found $\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as the best approximation, we can also calculate the least-squares error:

$$\mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

$$\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

so $\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\| = \sqrt{84}$.



Least-squares solutions

Sometimes, $\mathbf{A}^\top\mathbf{A}$ may be sensitive to round-off errors.
There is an alternative way to obtain least-squares solutions.

Theorem: If $\mathbf{A}$ is an $m \times n$ matrix with linearly independent columns, $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a unique least-squares solution $\forall\, \mathbf{b} \in \mathbb{R}^m$:

$$\hat{\mathbf{x}} = \mathbf{R}^{-1}\mathbf{Q}^\top\mathbf{b}$$

where $\mathbf{A} = \mathbf{Q}\mathbf{R}$ is a QR-factorisation of $\mathbf{A}$.

In practice, $\hat{\mathbf{x}}$ is obtained by solving

$$\mathbf{R}\hat{\mathbf{x}} = \mathbf{Q}^\top\mathbf{b}$$

which is straightforward since $\mathbf{R}$ is upper-triangular.

Least-squares solutions

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}$$

We know $\quad \mathbf{A} = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$

$\mathbf{Q}^\top \mathbf{b} = \begin{bmatrix} 1 \\ 2/\sqrt{12} \\ 2/\sqrt{6} \end{bmatrix}$ so $\left[\begin{array}{ccc|c} 1 & 3/2 & 1 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} & 2/\sqrt{6} \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$

thus $\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and, given that $\mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, the error is $\|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}\| = \sqrt{2}$.

Revision: eigenvectors and eigenvalues

Definition: An *eigenvector* of a square matrix $\mathbf{A}$ is a nonzero vector $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ.

A scalar λ is called an *eigenvalue* of $\mathbf{A}$ if there is a nontrivial solution $\mathbf{x}$ of $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

Such an $\mathbf{x}$ is an *eigenvector corresponding to* λ.

λ is an eigenvalue for $\mathbf{A}$ if and only there is a nontrivial solution to

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

λ can be found by solving a *characteristic equation*:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

$\mathrm{Nul}(\mathbf{A} - \lambda\mathbf{I})$ is called the *eigenspace* of $\mathbf{A}$ corresponding to λ.

Revision: eigenvectors and eigenvalues

- The eigenvalues of a triangular matrix are the entries on its main diagonal:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \qquad \Rightarrow \qquad \begin{cases} \lambda_1 = a_{11} \\ \lambda_2 = a_{22} \\ \lambda_3 = a_{33} \end{cases}$$

- 0 is an eigenvalue of $\mathbf{A}$ if and only if $\mathbf{A}$ is singular.

- If $\mathbf{v}_1, \ldots \mathbf{v}_p$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots \lambda_p$, then the set $\{\mathbf{v}_1, \ldots \mathbf{v}_p\}$ is linearly independent.

Revision: eigenvectors and eigenvalues

Example: for $\mathbf{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$ find bases for its eigenspaces.

Solution: $\det(\mathbf{A} - \lambda \mathbf{I}) = \cdots = (9-\lambda)(2-\lambda)(2-\lambda) = 0$
so $\lambda_1 = 9$ and $\lambda_2 = 2$; then we solve $(\mathbf{A} - \lambda \mathbf{I})\,\mathbf{x} = \mathbf{0}$

$$\mathbf{A} - 9\mathbf{I} = \begin{bmatrix} -5 & -1 & 6 \\ 2 & -8 & 6 \\ 2 & -1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} = x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{2} & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} = x_2 \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

So the bases are:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad \text{for } \lambda_1 \qquad \text{and} \qquad \left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\} \quad \text{for } \lambda_2$$

Matrix diagonalisation

- A square matrix $\mathbf{A}$ is *diagonalisable* if $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ where $\mathbf{D}$ is a diagonal matrix.

- An $n \times n$ matrix $\mathbf{A}$ is diagonalisable if and only if $\mathbf{A}$ has n linearly independent eigenvectors.

  The columns of $\mathbf{P}$ are then the eigenvectors of $\mathbf{A}$.

  The diagonal entries of $\mathbf{D}$ are the eigenvalues of $\mathbf{A}$ that correspond to the eigenvectors in $\mathbf{P}$.

- $\mathbf{A}$ is diagonalisable if and only if its eigenvectors form a basis of $\mathbb{R}^n$, which is called the *eigenvector basis*.

- An $n \times n$ matrix with n distinct eigenvalues is diagonalisable.

  Note: this is a sufficient, but not a necessary condition.

Matrix diagonalisation

Example: for $\mathbf{A} = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$, we have found the eigenvalues

$\lambda_1 = 9$ and $\lambda_2 = 2$, and three linearly independent eigenvectors

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (\text{for } \lambda_1) \qquad \text{and} \qquad \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \quad (\text{for } \lambda_2)$$

Then $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ where

$$\mathbf{P} = \begin{bmatrix} 1 & 1/2 & -3 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \mathbf{P}^{-1} = \frac{1}{7} \begin{bmatrix} 2 & -1 & 6 \\ -2 & 8 & -6 \\ -2 & 1 & 1 \end{bmatrix}$$

Orthogonal diagonalisation

Recall: a matrix $\mathbf{A}$ is symmetric if $\mathbf{A}^\top = \mathbf{A}$.

Theorem: If $\mathbf{A}$ is symmetric, then any two eigenvectors from different eigenspaces (corresponding to distinct eigenvalues) are orthogonal.

Proof: Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues $\lambda_1 \neq \lambda_2$. We must show that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

$$\begin{aligned}
\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^\top \mathbf{v}_2 = (\mathbf{A}\mathbf{v}_1)^\top \mathbf{v}_2 = (\mathbf{v}_1^\top \mathbf{A}^\top)\mathbf{v}_2 \\
&= (\mathbf{v}_1^\top \mathbf{A})\mathbf{v}_2 = \mathbf{v}_1{}^\top(\mathbf{A}\mathbf{v}_2) = \mathbf{v}_1{}^\top(\lambda_2 \mathbf{v}_2) \\
&= \lambda_2 \mathbf{v}_1^\top \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2.
\end{aligned}$$

Therefore $(\lambda_1 - \lambda_2)\, \mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$

But $\lambda_1 - \lambda_2 \neq 0$ hence $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

Orthogonal diagonalisation

Definition: A matrix $\mathbf{A}$ is *orthogonally diagonalisable* if there are an orthogonal matrix $\mathbf{P}$ $(\mathbf{P}^{-1} = \mathbf{P}^\top)$ and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^\top = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$$

Note: If $\mathbf{A}$ is orthogonally diagonalisable then

$$\mathbf{A}^\top = \left(\mathbf{P}\mathbf{D}\mathbf{P}^\top\right)^\top = \left(\mathbf{P}^\top\right)^\top \mathbf{D}^\top \mathbf{P}^\top = \mathbf{P}\mathbf{D}\mathbf{P}^\top = \mathbf{A}$$

and therefore $\mathbf{A}$ is symmetric. In fact:

Theorem: An $n \times n$ matrix $\mathbf{A}$ is orthogonally diagonalisable if and only if $\mathbf{A}$ is a symmetric matrix.

Orthogonal diagonalisation

Example: orthogonally diagonalise $\mathbf{A} = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$

Solution: The characteristic equation for this matrix is

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2).$$

Eigenvalues are $\lambda_{1,2} = 7$ (with multiplicity 2), and $\lambda_3 = -2$.

$$\mathbf{A} - 7\mathbf{I} = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \to \begin{bmatrix} 1 & 1/2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \left\{ \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\mathbf{A} + 2\mathbf{I} = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \left\{ \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix} \right\}$$

Orthogonal diagonalisation

The eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}$ are linearly independent but not orthogonal. Use the Gram-Schmidt process:

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

and then normalise the vectors for an orthonormal set:

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

The third eigenvector (for λ_3) is already orthogonal to this pair and we only need to normalise it:

$$\mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{u}_3 = \frac{2\mathbf{v}_3}{\|2\mathbf{v}_3\|} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

Orthogonal diagonalisation

So the orthonormal set of eigenvectors is

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

In this way, we have obtained $\mathbf{P}$ and $\mathbf{D}$:

$$\mathbf{P} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

and $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1} = \mathbf{P}\mathbf{D}\mathbf{P}^{\top}$.

Spectral theorem

The set of eigenvalues of a matrix $\mathbf{A}$ is called the *spectrum* of $\mathbf{A}$.

Theorem:

An $n \times n$ symmetric matrix $\mathbf{A}$ has the following properties:

(a) $\mathbf{A}$ has n real eigenvalues counting multiplicities;

(b) The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation;

(c) The eigenspaces are mutually orthogonal (the eigenvectors corresponding to different eigenvalues are orthogonal);

(d) $\mathbf{A}$ is orthogonally diagonalisable: $\mathbf{A} = \mathbf{PDP}^{-1} = \mathbf{PDP}^{\top}$.

Spectral decomposition

For an orthogonally diagonalisable matrix:

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^\top = \begin{bmatrix}\mathbf{u}_1 & \dots & \mathbf{u}_n\end{bmatrix} \begin{bmatrix}\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix} \begin{bmatrix}\mathbf{u}_1^\top \\ \vdots \\ \mathbf{u}_n^\top\end{bmatrix}$$

$$= \lambda_1 \mathbf{u}_1 \mathbf{u}_1^\top + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^\top + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^\top.$$

This is called a *spectral decomposition* of $\mathbf{A}$.

- Each decomposition term is an $n \times n$ matrix with rank 1.
- Each matrix $\mathbf{u}_j \mathbf{u}_j^\top$ is a projection matrix:

  for $\mathbf{x} \in \mathbb{R}^n$, vector $\mathbf{u}_j \mathbf{u}_j^\top \mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the subspace spanned by $\mathbf{u}_j$.

Spectral decomposition

Example: Spectral decomposition of matrix $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$:

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^\top = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Here, $\mathbf{P} = \left[\mathbf{u}_1, \, \mathbf{u}_2\right]$, then $\mathbf{A} = 8\mathbf{u}_1\mathbf{u}_1^\top + 3\mathbf{u}_2\mathbf{u}_2^\top$.

Verifying this decomposition:

$$\mathbf{u}_1\mathbf{u}_1^\top = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^\top = \begin{bmatrix} \frac{-1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

$$8\mathbf{u}_1\mathbf{u}_1^\top + 3\mathbf{u}_2\mathbf{u}_2^\top = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$$

Quadratic forms

Definition: A *quadratic form* on $\mathbb{R}^n$ is a function Q defined as

$$Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} \qquad : \qquad \mathbf{x} \in \mathbb{R}^n \qquad \mathbf{A} = \mathbf{A}^\top$$

where the $n \times n$ symmetric $\mathbf{A}$ is the *matrix of the quadratic form*.

Examples: (1) The simplest QF is: $\mathbf{x}^\top \mathbf{I}\mathbf{x} = \mathbf{x}^\top \mathbf{x} = \|\mathbf{x}\|^2$.

(2) Let $\mathbf{A} = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$, $\quad \mathbf{B} = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$, $\quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2$$

$$\begin{aligned} \mathbf{x}^\top \mathbf{B}\mathbf{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\ &= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2) = 3x_1^2 - 4x_1x_2 + 7x_2^2 \end{aligned}$$

Quadratic forms

Example: $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3 \qquad \mathbf{x} \in \mathbb{R}^3$

Let us write this quadratic form as $\mathbf{x}^\top A\mathbf{x}$:

The coefficients of x_1^2, x_2^2, x_3^2 provide the diagonal of $\mathbf{A}$.

Then, to make $\mathbf{A}$ symmetric we split the coefficients of x_ix_j between the i, j and j, i matrix elements:

$$Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Quadratic forms

Theorem (the principal axes theorem):

For a given quadratic form $\mathbf{x}^\top \mathbf{A}\mathbf{x}$ there is a change of basis (change of variable) $\mathbf{x} = \mathbf{P}\mathbf{y}$ that transforms it into a quadratic form $\mathbf{y}^\top \mathbf{D}\mathbf{y}$ with a diagonal matrix $\mathbf{D}$ (no cross-product terms).

Proof:

Since $\mathbf{A} = \mathbf{A}^\top$, it can be orthogonally diagonalised, $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^\top$. Introduce a change of basis $\mathbf{x} = \mathbf{P}\mathbf{y}$, so then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$. Then

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} = (\mathbf{P}\mathbf{y})^\top \mathbf{A}(\mathbf{P}\mathbf{y}) = \mathbf{y}^\top \mathbf{P}^\top (\mathbf{P}\mathbf{D}\mathbf{P}^\top)\mathbf{P}\mathbf{y} = \mathbf{y}^\top \mathbf{D}\mathbf{y}$$

and so the matrix in the quadratic form for $\mathbf{y}$ is diagonal.

Notes:

The columns of $\mathbf{P}$ are called the *principal axes* of the QF $\mathbf{x}^\top \mathbf{A}\mathbf{x}$.

Vector $\mathbf{y}$ gives coordinates of $\mathbf{x}$ relative to the principal axes, which form an orthonormal basis for $\mathbb{R}^n$.

Quadratic forms

Example: $Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2 \quad \Leftrightarrow \quad \mathbf{A} = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}.$

Let us find the principal axes and eliminate the cross terms.

The eigenvalues are $\lambda_1 = 3$, $\lambda_2 = -7$, and the unit eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

are orthogonal because $\mathbf{A}$ is symmetric and $\lambda_1 \neq \lambda_2$.

These vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are the principle axes of $Q(\mathbf{x})$.

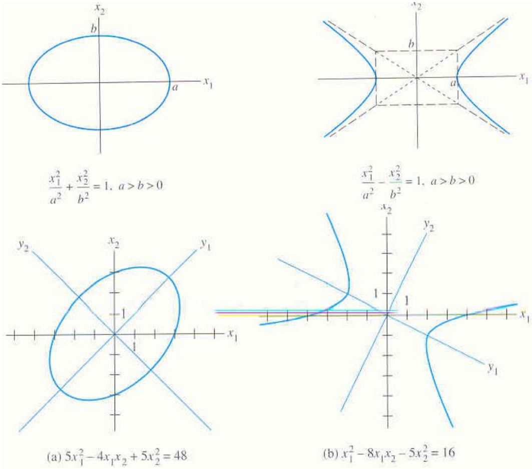
$$\mathbf{P} = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}.$$

The change of variable is $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x} = \mathbf{P}^\top \mathbf{x}$, and $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^\top$. So

$$x_1^2 - 8x_1x_2 - 5x_2^2 = \mathbf{x}^\top \mathbf{A}\mathbf{x} = \mathbf{y}^\top \mathbf{D}\mathbf{y} = 3y_1^2 - 7y_2^2$$

Quadratic forms

Example: $\mathbf{x}^\top \mathbf{A}\mathbf{x} = c$ for $\mathbf{x} \in \mathbb{R}^2$ and $c \in \mathbb{R}$
describes an ellipse, hyperbola, parabola, two intersecting lines, a single point, or no point at all.



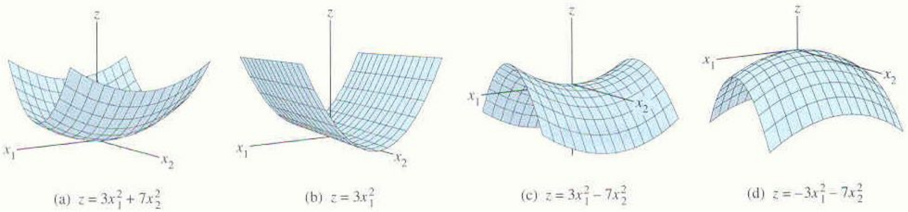
If $\mathbf{A}$ is diagonal then the graph is in the standard position.

$$\mathbf{A}_{(\mathrm{a})} = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$$

$$\mathbf{A}_{(\mathrm{b})} = \begin{bmatrix} 5 & -4 \\ -4 & -5 \end{bmatrix}$$

Quadratic forms

A few examples of $z = Q(\mathbf{x})$ for some typical cases ($\mathbf{x} \in \mathbb{R}^2$):



Definition: A quadratic form Q is

(a) *positive definite*, if $Q(\mathbf{x}) > 0 \ \forall \mathbf{x} \neq 0$ — all $\lambda > 0$

(b) *positive semidefinite*, if $Q(\mathbf{x}) \geqslant 0 \ \forall \mathbf{x}$

(d) *indefinite*, if $Q(\mathbf{x})$ takes positive and negative values

(c) *negative definite*, if $Q(\mathbf{x}) < 0 \ \forall \mathbf{x} \neq 0$ — all $\lambda < 0$

(e) *negative semidefinite*, if $Q(\mathbf{x}) \leqslant 0 \ \forall \mathbf{x}$

Singular value decomposition

- Orthogonal diagonalisation is a very useful tool however only symmetric matrices can be decomposed as $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$.
- However, a more general decomposition: $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{P}^{-1}$ (with $\mathbf{D}$ diagonal) is possible for any $m \times n$ matrix $\mathbf{A}$.
- If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}^\top\mathbf{A}$ is symmetric and can be orthogonally diagonalised.
- Let $\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ be the unit eigenvectors of $\mathbf{A}^\top\mathbf{A}$, and $\lambda_1, \ldots \lambda_n$ be the corresponding eigenvalues. Then

$$\|\mathbf{A}\mathbf{v}_i\|^2 = (\mathbf{A}\mathbf{v}_i)^\top\mathbf{A}\mathbf{v}_i = \mathbf{v}_i^\top\mathbf{A}^\top\mathbf{A}\mathbf{v}_i = \mathbf{v}_i^\top(\lambda_i\mathbf{v}_i) = \lambda_i,$$

  therefore all the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ are non-negative.
- We can always rearrange them in descending order so that

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

Singular value decomposition

Definition: The *singular values* of $\mathbf{A}$ are the square roots of the eigenvalues $\lambda_1, \ldots \lambda_n$ of $\mathbf{A}^\top \mathbf{A}$, arranged in the descending order:

$$\sigma_i = \sqrt{\lambda_i} \qquad \sigma_i \geqslant \sigma_{i+1}$$

Note: The singular values of $\mathbf{A}$ are the lengths of $\mathbf{A}\mathbf{v}_1, \ldots \mathbf{A}\mathbf{v}_n$.

Theorem: Let $\mathbf{A}$ be an $m \times n$ matrix with rank r. Then:

- there is an $m \times n$ matrix Σ with the first r diagonal entries being the singular values of $\mathbf{A}$: $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0$,
- there is an $m \times m$ orthogonal matrix $\mathbf{W}$,
- there is an $n \times n$ orthogonal matrix $\mathbf{U}$,

such that $\quad \mathbf{A} = \mathbf{W}\Sigma\mathbf{U}^\top$.

This decomposition is called a *singular value decomposition* of $\mathbf{A}$.

Singular value decomposition

Notes:

- The decomposition of $\mathbf{A}$ involves an $m \times n$ "quasi-diagonal" matrix $\Sigma = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, where $\mathbf{D}$ is an $r \times r$ diagonal matrix ($r \leqslant m$ & $r \leqslant n$).

  The second line in Σ contains $m - r$ rows.

  The second column in Σ contains $n - r$ columns.

- The matrices $\mathbf{U}$ and $\mathbf{W}$ in $\mathbf{A} = \mathbf{W}\Sigma\mathbf{U}^\top$ are not uniquely defined by $\mathbf{A}$ but the diagonal entries in Σ are uniquely determined (by the singular values of $\mathbf{A}$).

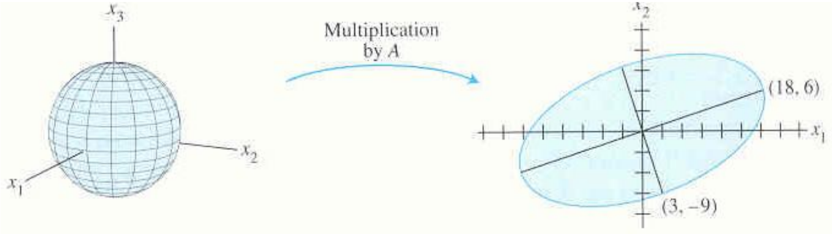
- The columns of $\mathbf{W}$ are called *left singular vectors* of $\mathbf{A}$ and the columns of $\mathbf{U}$ are called the *right singular vectors* of $\mathbf{A}$.

Singular value decomposition

Example: Construct a singular value decomposition of

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

Note: The transformation $\mathbf{x} \to \mathbf{A}\mathbf{x}$ maps a unit sphere $\{\mathbf{x} : \|\mathbf{x}\| = 1\}$ in $\mathbb{R}^3$ onto an ellipse in $\mathbb{R}^2$.



Example: Construct a singular value decomposition of

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

Step 1: Construct an orthogonal diagonalisation of $\mathbf{A}^\top\mathbf{A}$.

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

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

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

Then $$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}.$$

Singular value decomposition

Step 2: Construct Σ using the singular values of $\mathbf{A}$.

Given the found eigenvalues of $\mathbf{A}$, $\lambda_1 = 360$, $\lambda_2 = 90$, $\lambda_3 = 0$, 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 non-zero σ_i are the diagonal values of $\mathbf{D}$:

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

$$\Sigma = \left[\mathbf{D}\ \mathbf{0}\right] = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

Singular value decomposition

Step 3: Construct $\mathbf{W}$. When $\mathbf{A}$ has rank r the first r columns of $\mathbf{W}$ are normalised vectors obtained from $\mathbf{Av}_1, \ldots \mathbf{Av}_r$. $\mathbf{A}$ has two non-zero singular values so $\operatorname{rank} \mathbf{A} = 2$ and

$$\|\mathbf{Au}_1\| = \sigma_1, \quad \|\mathbf{Au}_2\| = \sigma_2.$$

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

$$\mathbf{w}_1 = \frac{1}{\sigma_1}\mathbf{Au}_1 = \frac{1}{6\sqrt{10}}\begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$
$$\mathbf{w}_2 = \frac{1}{\sigma_2}\mathbf{Au}_2 = \frac{1}{3\sqrt{10}}\begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}.$$

The set $\{\mathbf{w}_1, \mathbf{w}_2\}$ is already a basis for $\mathbb{R}^2$, and so

$$\mathbf{W} = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}$$

.

Singular value decomposition

Thus the singular value decomposition is

$$\mathbf{A} = \mathbf{W}\Sigma\mathbf{U}^\top$$

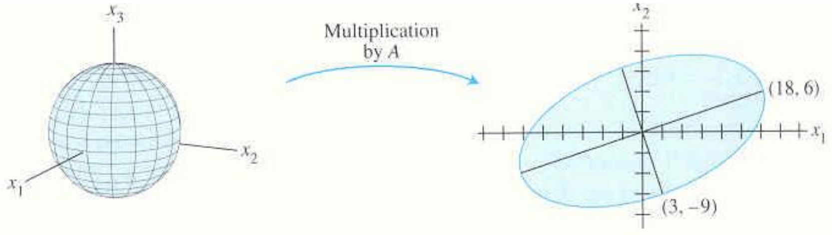
$$= \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{-2}{3} & \frac{-1}{3} & \frac{2}{3} \\ \frac{2}{3}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

Singular value decomposition

The first singular value of $\mathbf{A}$ is the maximum of $\|\mathbf{Ax}\|$ for all $\mathbf{x}$ with $\|\mathbf{x}\| = 1$; this is attained when $\mathbf{x} = \mathbf{u}_1$:

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

This is an ellipse point furthest from $\mathbf{0}$; the distance is $\sigma_1 = 6\sqrt{10}$:

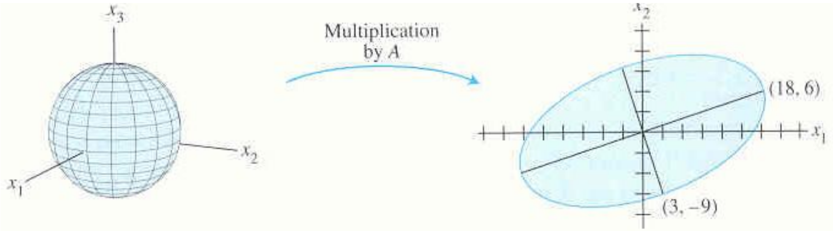


Singular value decomposition

The second singular value of $\mathbf{A}$ is the maximum of $\|\mathbf{Ax}\|$ over all unit vectors orthogonal to $\mathbf{u}_1$ and this is attained at $\mathbf{x} = \mathbf{u}_2$:

$$\mathbf{A}\mathbf{u}_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}.$$

This is an ellipse point on the minor axis (distance $\sigma_1 = 3\sqrt{10}$):



Singular value decomposition

$\mathbf{A} = \mathbf{W}\Sigma\mathbf{U}^\top$ can be rewritten as

$$\mathbf{A} = \begin{bmatrix}\mathbf{w}_1 & \dots & \mathbf{w}_m\end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 & 0 & \\ & \ddots & & 0 & \\ 0 & & \sigma_r & 0 & \dots \\ 0 & 0 & 0 & 0 & \\ & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix}\mathbf{v}_1^\top \\ \vdots \\ \mathbf{v}_n^\top\end{bmatrix}$$

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

Original matrix $\mathbf{A}$ involves $m \times n$ values to be stored, whereas this expansion requires $(m \times r + n \times r + r) = r(m + n + 1)$.

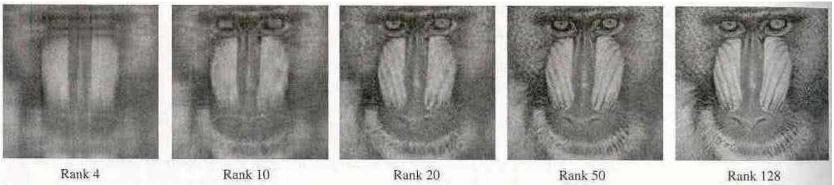
Singular value decomposition

Usually some of the singular values are very small so

$$\mathbf{A} \approx \mathbf{A}_k = \sigma_1 \mathbf{w}_1 \mathbf{u}_1^\top + \sigma_2 \mathbf{w}_2 \mathbf{u}_2^\top + \ldots + \sigma_k \mathbf{w}_k \mathbf{u}_k^\top,$$

where $k < r$ is the *rank of approximation*; quite often $k \ll r$.

In that case, the storage size is reduced to $k(m+n+1) \ll m \cdot n$.



Then SVD-based image compression / dimension reduction works.

Next lecture

See you next Wednesday for the final lecture

5 June 2019

Assignment 8 is due this week (on 29–31 May)

Assignment 9 will be an in-class pre-exam test

(during the tutorial sessions next week)