37233 Linear Algebra — revision

- Revision of the key topics
- Your questions?

Key topics

- Matrix subspaces
- Scalar product and orthogonality; projections; orthogonal basis
- Diagonalisation and SVD
- Quadratic forms

Null, column, row spaces

For an $m \times n$ matrix $\mathbf{A}$, the following subspaces are defined:

- *Null space*: all solutions to $\mathbf{A}\mathbf{x} = \mathbf{0}$:

$$\mathrm{Nul}\,\mathbf{A} = \{\mathbf{x} \,:\, \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

- *Column space*: all linear combinations of the columns of $\mathbf{A}$:

$$\mathrm{Col}\,\mathbf{A} = \mathrm{Span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots \mathbf{a}_n\}$$

$$\mathrm{Col}\,\mathbf{A} = \{\mathbf{b} \,:\, \mathbf{b} = \mathbf{A}\mathbf{x} \text{ for } \mathbf{x} \in \mathbb{R}^n\}$$

- *Row space*: all linear combinations of the rows of $\mathbf{A}$:

$$\mathrm{Row}\,\mathbf{A} = \mathrm{Col}\,\mathbf{A}^\mathsf{T}$$

For a given $m \times n$ matrix $\mathbf{A}$:

- $\operatorname{Nul}\mathbf{A} \subset \mathbb{R}^n$; $\quad \operatorname{Col}\mathbf{A} \subset \mathbb{R}^m$; $\quad \operatorname{Row}\mathbf{A} \subset \mathbb{R}^n$.
- $\operatorname{Col}\mathbf{A} = \operatorname{Span}\{\mathbf{a}_i\}$ and $\operatorname{Row}\mathbf{A} = \operatorname{Span}\{\mathbf{z}_i\}$ (here $\mathbf{Z} = \mathbf{A}^\mathsf{T}$) are explicitly defined. $\operatorname{Nul}\mathbf{A}$ is implicitly defined by $\mathbf{A}\mathbf{x} = \mathbf{0}$.
- There is no direct relation between $\operatorname{Nul}\mathbf{A}$ and a_{ij}
  There is a direct relation between $\operatorname{Col}\mathbf{A}$, $\operatorname{Row}\mathbf{A}$ and a_{ij}
- A vector of $\operatorname{Nul}\mathbf{A}$ is obtained by solving $\mathbf{A}\mathbf{x} = \mathbf{0}$.
  A vector of $\operatorname{Col}\mathbf{A}$ is obtained as a linear combination of $\{\mathbf{a}_i\}$, and of $\operatorname{Row}\mathbf{A}$ as a linear combination of $\{\mathbf{z}_i\}$.
- Checking if $\mathbf{v} \in \operatorname{Nul}\mathbf{A}$ is done by computing if $\mathbf{A}\mathbf{v} = \mathbf{0}$
  Checking if $\mathbf{v} \in \operatorname{Col}\mathbf{A}$ requires solving $\mathbf{A}\mathbf{x} = \mathbf{v}$
  Checking if $\mathbf{v} \in \operatorname{Row}\mathbf{A}$ requires solving $\mathbf{A}^\mathsf{T}\mathbf{x} = \mathbf{v}$
- $\operatorname{Nul}\mathbf{A} = \{\mathbf{0}\}$ if and only if $\mathbf{A}\mathbf{x} = \mathbf{0}$ only for $\mathbf{x} = \mathbf{0}$
  $\operatorname{Col}\mathbf{A} = \mathbb{R}^m$ if and only if $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a solution $\forall\, \mathbf{b} \in \mathbb{R}^m$
  $\operatorname{Row}\mathbf{A} = \mathbb{R}^n$ if and only if $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a solution $\forall\, \mathbf{b} \in \mathbb{R}^m$
- $\operatorname{Nul}\mathbf{A} = \left(\operatorname{Row}\mathbf{A}\right)^\perp$ and $(\operatorname{Col}\mathbf{A})^\perp = \operatorname{Nul}\left(\mathbf{A}^\mathsf{T}\right)$

Bases and dimensions for $\operatorname{Nul}\mathbf{A}$, $\operatorname{Col}\mathbf{A}$, and $\operatorname{Row}\mathbf{A}$

- A basis for $\operatorname{Nul}\mathbf{A}$ is provided by linearly independent vectors spanning the solution of a homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{0}$.
- A basis for $\operatorname{Col}\mathbf{A}$ is formed by pivot columns of a $\mathbf{A}$.

  Warning: It is important that the pivot columns of $\mathbf{A}$ itself, and not those of the REF form, are a basis for $\operatorname{Col}\mathbf{A}$.
- A basis for $\operatorname{Row}\mathbf{A}$ is formed by pivot rows of a $\mathbf{A}$.

  Note: The row space of the REF of $\mathbf{A}$ is the same space.

By proceeding with solving $\mathbf{A}\mathbf{x} = \mathbf{0}$ we can establish:

- $\dim(\operatorname{Nul}\mathbf{A})$ is the number of free variables.
- $\dim(\operatorname{Col}\mathbf{A})$ is the number of pivot columns.
- $\dim(\operatorname{Row}\mathbf{A})$ is the number of pivot rows.
- $\operatorname{rank}\mathbf{A} = \dim(\operatorname{Col}\mathbf{A}) = \dim(\operatorname{Row}\mathbf{A})$
- $\operatorname{rank}\mathbf{A} + \dim(\operatorname{Nul}\mathbf{A}) = n$ (given $m \times n$ matrix $\mathbf{A}$)

The invertible matrix theorem (summary of results)

Statements equivalent to $\mathbf{A}$ being an $n \times n$ invertible matrix:

- There is an $n \times n$ matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$
- $\mathbf{A}$ has n pivot positions in the REF form
- $\mathbf{A}\mathbf{x} = \mathbf{0}$ has only the trivial solution
- The columns (rows) of $\mathbf{A}$ form a linearly independent set
- The columns (rows) of $\mathbf{A}$ span $\mathbb{R}^n$
- The columns (rows) of $\mathbf{A}$ form a basis of $\mathbb{R}^n$
- $\widehat{T}: \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ is one-to-one
- $\widehat{T}: \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$
- $\mathrm{Col}\,\mathbf{A} = \mathrm{Row}\,\mathbf{A} = \mathbb{R}^n$
- $\mathrm{Nul}\,\mathbf{A} = \{\mathbf{0}\}$ and $\dim(\mathrm{Nul}\,\mathbf{A}) = 0$
- $\dim(\mathrm{Col}\,\mathbf{A}) = \dim(\mathrm{Row}\,\mathbf{A}) = n$
- $\mathrm{rank}\,\mathbf{A} = n$
- The eigenvalues of $\mathbf{A}$ are non-zero

Scalar product, orthogonality, norm

For $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^n$, the product $\mathbf{v} \cdot \mathbf{w} \equiv \mathbf{v}^\intercal \mathbf{w} = \sum_{i=1}^{n} v_i w_i$

is called the *scalar product* (or inner product, or dot product).

$\forall \{\mathbf{v}, \mathbf{w}, \mathbf{x}\} \in \mathbb{R}^n$:

- $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$
- $(\mathbf{v} + \mathbf{w}) \cdot \mathbf{x} = \mathbf{v} \cdot \mathbf{x} + \mathbf{w} \cdot \mathbf{x}$
- $(c\,\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (c\,\mathbf{w}) = c\,(\mathbf{v} \cdot \mathbf{w})$
- $\mathbf{v} \cdot \mathbf{v} \geqslant 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$

Vectors are *orthogonal* ($\mathbf{v} \perp \mathbf{w}$) if and only if $\mathbf{v} \cdot \mathbf{w} = 0$.

The length, or *norm*, of $\mathbf{v}$ is: $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\mathbf{v}^\intercal \mathbf{v}}$.

A *unit vector* has a unit norm (length): $\mathbf{u} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$

Orthogonal sets

Definition: A set of vectors $\{\mathbf{u}_1, \dots \mathbf{u}_p\}$ in $\mathbb{R}^n$ is called an *orthogonal set* if each pair of vectors from the set is orthogonal:

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0 \qquad \forall \ i \neq j$$

Definition: An *orthogonal basis* for a subspace V of $\mathbb{R}^n$ is such a basis for V which is an orthogonal set.

Coordinates with respect to an orthogonal basis are easily found:

$$\mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_p\mathbf{u}_p \qquad \text{has the weights} \qquad c_i = \frac{\mathbf{x} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}$$

Definition: A set $\{\mathbf{u}_1, \dots \mathbf{u}_p\}$ is called an *orthonormal set* if it is an orthogonal set of unit vectors.

The standard basis $\{\mathbf{e}_1, \dots \mathbf{e}_n\}$ for $\mathbb{R}^n$ is an orthonormal set.

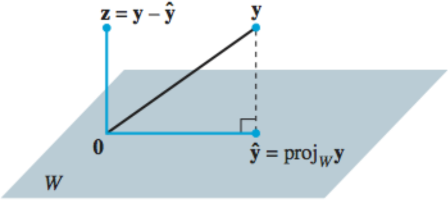
Orthogonal decomposition

Let W be a subspace of $\mathbb{R}^n$.

Then $\forall\, \mathbf{y} \in \mathbb{R}^n$ there is a unique decomposition

$$\mathbf{y} = \check{\mathbf{y}} + \mathbf{z},$$

where $\check{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$.



If (and only if) $\{\mathbf{u}_1, \ldots \mathbf{u}_p\}$ is any orthogonal basis in W,

$$\check{\mathbf{y}} = \sum_{i=1}^{p} \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}\, \mathbf{u}_i \qquad \text{and} \qquad \mathbf{z} = \mathbf{y} - \check{\mathbf{y}}.$$

$\check{\mathbf{y}} \equiv \mathrm{proj}_W\, \mathbf{y}$ is called the *orthogonal projection* of $\mathbf{y}$ onto W.

For an *orthonormal* set: $\check{\mathbf{y}} = \mathbf{U}\mathbf{U}^\top \mathbf{y}$ (where $\mathbf{U} = [\mathbf{u}_1\, \mathbf{u}_2\, \ldots\, \mathbf{u}_p]$).

Gram-Schmidt process

Given a basis $\mathbf{x}_1, \dots \mathbf{x}_p$ for a subspace W of $\mathbb{R}^n$, define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\, \mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\, \mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\, \mathbf{v}_2$$

$$\dots$$

$$\mathbf{v}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\, \mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\, \mathbf{v}_2 - \dots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\, \mathbf{v}_{p-1}$$

Then $\{\mathbf{v}_1, \dots \mathbf{v}_p\}$ is an orthogonal basis for W.

In addition, $\mathrm{Span}\{\mathbf{v}_1, \dots \mathbf{v}_k\} = \mathrm{Span}\{\mathbf{x}_1, \dots \mathbf{x}_k\}$ for $1 \leqslant k \leqslant p$.

Orthonormal basis is then obtained by normalising $\mathbf{v}_i$ to unit vectors.

Orthogonal matrices

Definition: An *orthogonal matrix* $\mathbf{U}$ is a square invertible matrix such that $\mathbf{U}^{-1} = \mathbf{U}^{\mathsf{T}}$.

Equivalent properties:

- $\mathbf{U}$ has orthonormal columns and orthonormal rows:

$$\sum_{k=1}^{n} u_{ki} u_{kj} = \delta_{ij} \qquad \text{and} \qquad \sum_{k=1}^{n} u_{ik} u_{jk} = \delta_{ij}$$

- Given two different orthonormal bases in $\mathbb{R}^n$, the change of basis matrices between such bases are orthogonal matrices.

Note: For an orthogonal matrix $\mathbf{U}$: $|\det \mathbf{U}| = 1$

Matrix diagonalisation

If $\mathbf{A}$ has n linearly independent eigenvectors, $\mathbf{A} = \mathbf{PDP}^{-1}$ where

- diagonal entries of the diagonal matrix $\mathbf{D}$ are the eigenvalues
- columns of $\mathbf{P}$ are the corresponding eigenvectors of $\mathbf{A}$

A symmetric matrix $\mathbf{A}$ is *orthogonally diagonalisable* as

$$\mathbf{A} = \mathbf{PDP}^{\top} = \mathbf{PDP}^{-1}$$

with $\mathbf{P}$ orthogonal $(\mathbf{P}^{-1} = \mathbf{P}^{\top})$ and $\mathbf{D}$ diagonal matrices.

Diagonalisation procedure:

1. Find the eigenvalues
2. Find the corresponding eigenvectors
3. Form matrices $\mathbf{D}$ and $\mathbf{P}$ accordingly

Singular value decomposition

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

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

SVD: for an $m \times n$ matrix $\mathbf{A}$ with rank r, construct:

- $m \times n$ matrix Σ with the first r diagonal entries being the singular values of $\mathbf{A}$: $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0$; then zeros
- $n \times n$ orthogonal matrix $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{bmatrix}$ (with $\mathbf{u}_i$ being the normalised eigenvectors of $\mathbf{A}^\mathsf{T}\mathbf{A}$)
- $m \times m$ orthogonal matrix $\mathbf{W}$: $\mathbf{w}_i = \dfrac{\mathbf{A}\mathbf{u}_i}{\sigma_i}$ for $1 \leqslant i \leqslant r$, extended to an orthonormal basis of $\mathbb{R}^m$ for $r < i \leqslant m$

Then $\mathbf{A} = \mathbf{W}\Sigma\mathbf{U}^\mathsf{T}$ is a *singular value decomposition* of $\mathbf{A}$.

Quadratic forms

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

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

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

The principal axes theorem:

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

Procedure:

1. Orthogonally diagonalise $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{\mathsf{T}}$
2. The required change of variables is $\mathbf{x} = \mathbf{P}\mathbf{y}$ and $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$
3. The columns of $\mathbf{P}$ are the *principal axes* of $\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$

Notes on the final test

- 110 minutes writing time plus 5 minutes technical time
- 3 problems covering the following main topics:
  - Matrix subspaces
  - Orthogonality, projections, orthogonal basis
  - Diagonalisation; spectral decomposition
  - Singular value decomposition
  - Quadratic forms
- Previous subject topics are still relevant (as solutions tools)

Advice for solving test problems

- Identify and attack the easiest problems first
- If stuck, change to the next problem and return at the end
- Keep to explicit radicals and rational fractions
  (e.g. $\dfrac{5}{\sqrt{3}}$, not "2.88675")
- If the numbers are getting really uncomfortable,
  something is probably wrong
- Articulate your solutions as much as practical
  (explain what you are doing)
- Check back the final answer where possible

Good luck!

Final class test

at the tutorials this week

Thank you for studying linear algebra, and good luck!