Lecture 15
Symmetric matrices, quadratic forms, matrix norm, and SVD

• eigenvectors of symmetric matrices

• quadratic forms

• inequalities for quadratic forms

• positive semidefinite matrices

• norm of a matrix

• singular value decomposition
Eigenvalues of symmetric matrices

suppose \( A \in \mathbb{R}^{n \times n} \) is symmetric, \( i.e., A = A^T \)

**fact:** the eigenvalues of \( A \) are real

to see this, suppose \( Av = \lambda v, v \neq 0, v \in \mathbb{C}^n \)

then

\[
\overline{v}^T Av = \overline{v}^T (Av) = \lambda \overline{v}^T v = \lambda \sum_{i=1}^{n} |v_i|^2
\]

but also

\[
\overline{v}^T Av = (Av)^T v = (\lambda v)^T v = \overline{\lambda} \sum_{i=1}^{n} |v_i|^2
\]

so we have \( \lambda = \overline{\lambda}, i.e., \lambda \in \mathbb{R} \) (hence, can assume \( v \in \mathbb{R}^n \))

Symmetric matrices, quadratic forms, matrix norm, and SVD
Eigenvectors of symmetric matrices

**Fact:** there is a set of orthonormal eigenvectors of $A$, i.e., $q_1, \ldots, q_n$ s.t. $Aq_i = \lambda_i q_i$, $q_i^T q_j = \delta_{ij}$

in matrix form: there is an orthogonal $Q$ s.t.

$$Q^{-1}AQ = Q^T AQ = \Lambda$$

hence we can express $A$ as

$$A = Q\Lambda Q^T = \sum_{i=1}^{n} \lambda_i q_i q_i^T$$

in particular, $q_i$ are both left and right eigenvectors
Interpretations

\[ A = Q\Lambda Q^T \]

linear mapping \( y = Ax \) can be decomposed as

- resolve into \( q_i \) coordinates
- scale coordinates by \( \lambda_i \)
- reconstitute with basis \( q_i \)
or, geometrically,

- rotate by $Q^T$
- diagonal real scale (‘dilation’) by $\Lambda$
- rotate back by $Q$

**decomposition**

$$A = \sum_{i=1}^{n} \lambda_i q_i q_i^T$$

expresses $A$ as linear combination of 1-dimensional projections
example:

\[
A = \begin{bmatrix}
-1/2 & 3/2 \\
3/2 & -1/2
\end{bmatrix}
\]

\[
= \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right)^T
\]
**proof (case of** \( \lambda_i \) **distinct)**

since \( \lambda_i \) distinct, can find \( v_1, \ldots, v_n \), a set of linearly independent eigenvectors of \( A \):

\[
Av_i = \lambda_i v_i, \quad \|v_i\| = 1
\]

then we have

\[
v_i^T (Av_j) = \lambda_j v_i^T v_j = (Av_i)^T v_j = \lambda_i v_i^T v_j
\]

so \((\lambda_i - \lambda_j)v_i^T v_j = 0\)

for \( i \neq j, \lambda_i \neq \lambda_j \), hence \( v_i^T v_j = 0 \)

• in this case we can say: eigenvectors are orthogonal

• in general case (\( \lambda_i \) not distinct) we must say: eigenvectors can be chosen to be orthogonal
Example: RC circuit

\[ c_k \dot{v}_k = -i_k, \quad i = Gv \]

\[ G = G^T \in \mathbb{R}^{n \times n} \] is conductance matrix of resistive circuit

thus \( \dot{v} = -C^{-1}Gv \) where \( C = \text{diag}(c_1, \ldots, c_n) \)

note \(-C^{-1}G\) is not symmetric
use state \( x_i = \sqrt{c_i}v_i \), so

\[
\dot{x} = C^{1/2}\dot{v} = -C^{-1/2}GC^{-1/2}x
\]

where \( C^{1/2} = \text{diag}(\sqrt{c_1}, \ldots, \sqrt{c_n}) \)

we conclude:

- eigenvalues \( \lambda_1, \ldots, \lambda_n \) of \(-C^{-1/2}GC^{-1/2}\) (hence, \(-C^{-1}G\)) are real
- eigenvectors \( q_i \) (in \( x_i \) coordinates) can be chosen orthogonal
- eigenvectors in voltage coordinates, \( s_i = C^{-1/2}q_i \), satisfy

\[
-C^{-1}Gs_i = \lambda_is_i, \quad s_i^TCs_i = \delta_{ij}
\]
Quadratic forms

a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(x) = x^T Ax = \sum_{i,j=1}^{n} A_{ij} x_i x_j$$

is called a quadratic form

in a quadratic form we may as well assume $A = A^T$ since

$$x^T Ax = x^T ((A + A^T)/2)x$$

$((A + A^T)/2$ is called the symmetric part of $A$)

**uniqueness:** if $x^T Ax = x^T Bx$ for all $x \in \mathbb{R}^n$ and $A = A^T$, $B = B^T$, then $A = B$
Examples

- \( \|Bx\|^2 = x^T B^T B x \)
- \( \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 \)
- \( \|F x\|^2 - \|G x\|^2 \)

sets defined by quadratic forms:

- \( \{ x \mid f(x) = a \} \) is called a \textit{quadratic surface}
- \( \{ x \mid f(x) \leq a \} \) is called a \textit{quadratic region}
Inequalities for quadratic forms

Suppose $A = A^T$, $A = Q\Lambda Q^T$ with eigenvalues sorted so $\lambda_1 \geq \cdots \geq \lambda_n$

\[
x^T Ax = x^T Q\Lambda Q^T x
= (Q^T x)^T \Lambda (Q^T x)
= \sum_{i=1}^{n} \lambda_i (q_i^T x)^2
\leq \lambda_1 \sum_{i=1}^{n} (q_i^T x)^2
= \lambda_1 \|x\|^2
\]

I.e., we have $x^T Ax \leq \lambda_1 x^T x$
similar argument shows $x^T Ax \geq \lambda_n \|x\|^2$, so we have

$$\lambda_n x^T x \leq x^T Ax \leq \lambda_1 x^T x$$

sometimes $\lambda_1$ is called $\lambda_{\text{max}}$, $\lambda_n$ is called $\lambda_{\text{min}}$

note also that

$$q_1^T A q_1 = \lambda_1 \|q_1\|^2, \quad q_n^T A q_n = \lambda_n \|q_n\|^2,$$

so the inequalities are tight
Positive semidefinite and positive definite matrices

suppose $A = A^T \in \mathbb{R}^{n \times n}$

we say $A$ is **positive semidefinite** if $x^T Ax \geq 0$ for all $x$

• denoted $A \geq 0$ (and sometimes $A \succeq 0$)
• $A \geq 0$ if and only if $\lambda_{\text{min}}(A) \geq 0$, i.e., all eigenvalues are nonnegative
• **not** the same as $A_{ij} \geq 0$ for all $i, j$

we say $A$ is **positive definite** if $x^T Ax > 0$ for all $x \neq 0$

• denoted $A > 0$
• $A > 0$ if and only if $\lambda_{\text{min}}(A) > 0$, i.e., all eigenvalues are positive
Matrix inequalities

• we say $A$ is *negative semidefinite* if $-A \geq 0$

• we say $A$ is *negative definite* if $-A > 0$

• otherwise, we say $A$ is *indefinite*

**matrix inequality:** if $B = B^T \in \mathbb{R}^n$ we say $A \geq B$ if $A - B \geq 0$, $A < B$ if $B - A > 0$, etc.

for example:

• $A \geq 0$ means $A$ is positive semidefinite

• $A > B$ means $x^T Ax > x^T B x$ for all $x \neq 0$
many properties that you’d guess hold actually do, e.g.,

- if $A \geq B$ and $C \geq D$, then $A + C \geq B + D$
- if $B \leq 0$ then $A + B \leq A$
- if $A \geq 0$ and $\alpha \geq 0$, then $\alpha A \geq 0$
- $A^2 \geq 0$
- if $A > 0$, then $A^{-1} > 0$

matrix inequality is only a partial order: we can have

$$A \not \succeq B, \quad B \not \succeq A$$

(such matrices are called incomparable)
Ellipsoids

if $A = A^T > 0$, the set

$$\mathcal{E} = \{ x \mid x^T A x \leq 1 \}$$

is an ellipsoid in $\mathbb{R}^n$, centered at 0.
semi-axes are given by \( s_i = \lambda_i^{-1/2} q_i \), i.e.:

- eigenvectors determine directions of semi-axes
- eigenvalues determine lengths of semi-axes

note:

- in direction \( q_1 \), \( x^T Ax \) is large, hence ellipsoid is thin in direction \( q_1 \)
- in direction \( q_n \), \( x^T Ax \) is small, hence ellipsoid is fat in direction \( q_n \)
- \( \sqrt{\lambda_{\text{max}}/\lambda_{\text{min}}} \) gives maximum eccentricity

if \( \tilde{\mathcal{E}} = \{ x \mid x^T B x \leq 1 \} \), where \( B > 0 \), then \( \mathcal{E} \subseteq \tilde{\mathcal{E}} \iff A \geq B \)
Gain of a matrix in a direction

suppose $A \in \mathbb{R}^{m \times n}$ (not necessarily square or symmetric)

for $x \in \mathbb{R}^n$, $\|Ax\|/\|x\|$ gives the *amplification factor* or *gain* of $A$ in the direction $x$

obviously, gain varies with direction of input $x$

questions:

• what is maximum gain of $A$
  (and corresponding maximum gain direction)?

• what is minimum gain of $A$
  (and corresponding minimum gain direction)?

• how does gain of $A$ vary with direction?
Matrix norm

the maximum gain

$$\max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

is called the matrix norm or spectral norm of $A$ and is denoted $\|A\|$

$$\max_{x \neq 0} \frac{\|Ax\|^2}{\|x\|^2} = \max_{x \neq 0} \frac{x^T A^T A x}{\|x\|^2} = \lambda_{\text{max}}(A^T A)$$

so we have $\|A\| = \sqrt{\lambda_{\text{max}}(A^T A)}$

similarly the minimum gain is given by

$$\min_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\lambda_{\text{min}}(A^T A)}$$
note that

- $A^T A \in \mathbb{R}^{n \times n}$ is symmetric and $A^T A \geq 0$ so $\lambda_{\text{min}}$, $\lambda_{\text{max}} \geq 0$

- ‘max gain’ input direction is $x = q_1$, eigenvector of $A^T A$ associated with $\lambda_{\text{max}}$

- ‘min gain’ input direction is $x = q_n$, eigenvector of $A^T A$ associated with $\lambda_{\text{min}}$
example: \( A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \)

\[
A^T A = \begin{bmatrix} 35 & 44 \\ 44 & 56 \end{bmatrix} = \begin{bmatrix} 0.620 & -0.785 \\ 0.785 & 0.620 \end{bmatrix} \begin{bmatrix} 90.7 & 0 \\ 0 & 0.265 \end{bmatrix} \begin{bmatrix} 0.620 & -0.785 \\ 0.785 & 0.620 \end{bmatrix}^T
\]

then \( \| A \| = \sqrt{\lambda_{\text{max}}(A^T A)} = 9.53: \)

\[
\| \begin{bmatrix} 0.620 \\ 0.785 \end{bmatrix} \| = 1, \quad \| A \begin{bmatrix} 0.620 \\ 0.785 \end{bmatrix} \| = \left\| \begin{bmatrix} 2.19 \\ 5.00 \\ 7.81 \end{bmatrix} \right\| = 9.53
\]
min gain is \( \sqrt{\lambda_{\text{min}}(A^T A)} = 0.514 \): 

\[
\left\| \begin{bmatrix} -0.785 \\ 0.620 \end{bmatrix} \right\| = 1, \quad \left\| A \begin{bmatrix} -0.785 \\ 0.620 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 0.45 \\ 0.12 \\ -0.21 \end{bmatrix} \right\| = 0.514
\]

for all \( x \neq 0 \), we have

\[
0.514 \leq \frac{\|Ax\|}{\|x\|} \leq 9.53
\]
Properties of matrix norm

• consistent with vector norm: matrix norm of $a \in \mathbb{R}^{n \times 1}$ is $\sqrt{\lambda_{\text{max}}(a^T a)} = \sqrt{a^T a}$

• for any $x$, $\|Ax\| \leq \|A\|\|x\|$

• scaling: $\|aA\| = |a|\|A\|$

• triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$

• definiteness: $\|A\| = 0 \iff A = 0$

• norm of product: $\|AB\| \leq \|A\|\|B\|$

Symmetric matrices, quadratic forms, matrix norm, and SVD
Singular value decomposition

more complete picture of gain properties of \( A \) given by *singular value decomposition* (SVD) of \( A \):

\[
A = U \Sigma V^T
\]

where

- \( A \in \mathbb{R}^{m \times n}, \text{Rank}(A) = r \)
- \( U \in \mathbb{R}^{m \times r}, U^T U = I \)
- \( V \in \mathbb{R}^{n \times r}, V^T V = I \)
- \( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_r), \text{where } \sigma_1 \geq \cdots \geq \sigma_r > 0 \)
with $U = [u_1 \cdots u_r]$, $V = [v_1 \cdots v_r]$, 

$$A = U \Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

- $\sigma_i$ are the (nonzero) singular values of $A$
- $v_i$ are the right or input singular vectors of $A$
- $u_i$ are the left or output singular vectors of $A$
\[
A^T A = (U \Sigma V^T)^T (U \Sigma V^T) = V \Sigma^2 V^T
\]

hence:

• \(v_i\) are eigenvectors of \(A^T A\) (corresponding to nonzero eigenvalues)

• \(\sigma_i = \sqrt{\lambda_i(A^T A)}\) (and \(\lambda_i(A^T A) = 0\) for \(i > r\))

• \(\|A\| = \sigma_1\)
similarly,

\[ AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma^2 U^T \]

hence:

- \( u_i \) are eigenvectors of \( AA^T \) (corresponding to nonzero eigenvalues)
- \( \sigma_i = \sqrt{\lambda_i(AA^T)} \) (and \( \lambda_i(AA^T) = 0 \) for \( i > r \))

- \( u_1, \ldots u_r \) are orthonormal basis for \( \text{range}(A) \)
- \( v_1, \ldots v_r \) are orthonormal basis for \( \mathcal{N}(A)^\perp \)
Interpretations

\[ A = U \Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T \]

linear mapping \( y = Ax \) can be decomposed as

- compute coefficients of \( x \) along input directions \( v_1, \ldots, v_r \)
- scale coefficients by \( \sigma_i \)
- reconstitute along output directions \( u_1, \ldots, u_r \)

difference with eigenvalue decomposition for symmetric \( A \): input and output directions are different
• $v_1$ is most sensitive (highest gain) input direction

• $u_1$ is highest gain output direction

• $Av_1 = \sigma_1 u_1$
SVD gives clearer picture of gain as function of input/output directions

**example:** consider $A \in \mathbb{R}^{4 \times 4}$ with $\Sigma = \text{diag}(10, 7, 0.1, 0.05)$

- input components along directions $v_1$ and $v_2$ are amplified (by about 10) and come out mostly along plane spanned by $u_1, u_2$

- input components along directions $v_3$ and $v_4$ are attenuated (by about 10)

- $\|Ax\|/\|x\|$ can range between 10 and 0.05

- $A$ is nonsingular

- for some applications you might say $A$ is *effectively* rank 2
\textbf{example:} \( A \in \mathbb{R}^{2\times 2}, \) with \( \sigma_1 = 1, \sigma_2 = 0.5 \)

- resolve \( x \) along \( v_1, v_2: \) \( v_1^T x = 0.5, \ v_2^T x = 0.6, \) \( i.e., \) \( x = 0.5v_1 + 0.6v_2 \)

- now form \( Ax = (v_1^T x)\sigma_1 u_1 + (v_2^T x)\sigma_2 u_2 = (0.5)(1)u_1 + (0.6)(0.5)u_2 \)