11. Quadratic forms and ellipsoids

- Quadratic forms
- Orthogonal decomposition
- Positive definite matrices
- Ellipsoids

Laurent Lessard (www.laurentlessard.com)

Quadratic forms

- **Linear functions:** sum of terms of the form $c_i x_i$ where the c_i are parameters and x_i are variables. General form:

$$c_1 x_1 + \cdots + c_n x_n = c^\mathsf{T} x$$

- **Quadratic functions:** sum of terms of the form $q_{ij} x_i x_j$ where q_{ij} are parameters and x_i are variables. General form:

$$q_{11} x_1^2 + q_{12} x_1 x_2 + \cdots + q_{nn} x_n^2 \qquad (n^2 \text{ terms})$$

$$= \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}^\mathsf{T} \begin{bmatrix} q_{11} & \dots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \dots & q_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x^\mathsf{T} Q x$$

Quadratic forms

Example: $4x^2 + 6xy - 2yz + y^2 - z^2$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^\mathsf{T} \begin{bmatrix} 4 & 6 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

In general: $$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^\mathsf{T} \begin{bmatrix} 4 & p_2 & q_2 \\ p_1 & 1 & r_2 \\ q_1 & r_1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \begin{cases} p_1 + p_2 = 6 \\ q_1 + q_2 = 0 \\ r_1 + r_2 = -2 \end{cases}$$

Symmetric: $$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^\mathsf{T} \begin{bmatrix} 4 & 3 & 0 \\ 3 & 1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Quadratic forms

Any quadratic function $f(x_1, \ldots, x_n)$ can be written in the form $x^\mathsf{T} Q x$ where Q is a symmetric matrix ($Q = Q^\mathsf{T}$).

Proof: Suppose $f(x_1, \ldots, x_n) = x^\mathsf{T} R x$ where R is *not* symmetric. Since it is a scalar, we can take the transpose:

$$x^\mathsf{T} R x = \left(x^\mathsf{T} R x\right)^\mathsf{T} = x^\mathsf{T} R^\mathsf{T} x$$

Therefore:

$$x^\mathsf{T} R x = \frac{1}{2}\left(x^\mathsf{T} R x + x^\mathsf{T} R^\mathsf{T} x\right) = x^\mathsf{T} \tfrac{1}{2}(R + R^\mathsf{T}) x$$

So we're done, because $\frac{1}{2}(R + R^\mathsf{T})$ is symmetric!

Orthogonal decomposition

Theorem. Every real symmetric matrix $Q = Q^\mathsf{T} \in \mathbb{R}^{n \times n}$ can be decomposed into a product:

$$Q = U \Lambda U^\mathsf{T}$$

where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ is a real diagonal matrix, and $U \in \mathbb{R}^{n \times n}$ is an orthogonal matrix. i.e. it satisfies $U^\mathsf{T} U = I$.

This is a useful decomposition because orthogonal matrices have very nice properties...

Orthogonal matrices

A matrix U is orthogonal if $U^\mathsf{T} U = I$.

- If the columns are $U = \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}$, then we have:

$$U^\mathsf{T} U = \begin{bmatrix} u_1^\mathsf{T} u_1 & \cdots & u_1^\mathsf{T} u_m \\ \vdots & \ddots & \vdots \\ u_m^\mathsf{T} u_1 & \cdots & u_m^\mathsf{T} u_m \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

  Columns of U are mutually orthogonal: $u_i^\mathsf{T} u_j = 0$ if $i \neq j$.

- If U is square, $U^{-1} = U^\mathsf{T}$, and U^T is also orthogonal.

Orthogonal matrices

- columns can be rearranged and the factorization stays valid.

$$
\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} u_1^\mathsf{T} \\ u_2^\mathsf{T} \\ u_3^\mathsf{T} \end{bmatrix}
$$

$$
= \lambda_1 u_1 u_1^\mathsf{T} + \lambda_2 u_2 u_2^\mathsf{T} + \lambda_3 u_3 u_3^\mathsf{T}
$$

$$
= \begin{bmatrix} u_1 & u_3 & u_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_3 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} u_1^\mathsf{T} \\ u_3^\mathsf{T} \\ u_2^\mathsf{T} \end{bmatrix}
$$

Orthogonal matrices

- Orthogonal matrices preserve angle and (2-norm) distance:

$$(Ux)^\mathsf{T}(Uy) = x^\mathsf{T}(U^\mathsf{T}U)y = x^\mathsf{T}y$$

In particular, we have $\|Uz\| = \|z\|$ for any z.

- If $Q = U\Lambda U^\mathsf{T}$, then multiply by u_i:

$$Qu_i = \begin{bmatrix} u_1^\mathsf{T} \\ \vdots \\ u_n^\mathsf{T} \end{bmatrix}^\mathsf{T} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} u_1^\mathsf{T} \\ \vdots \\ u_n^\mathsf{T} \end{bmatrix} u_i = \lambda_i u_i$$

So multiplication by Q simply scales each u_i by λ_i. In other words: (λ_i, u_i) are the eigenvalue-eigenvector pairs of Q.

Orthogonal matrix example

Rotation matrices are orthgonal:

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

We can verify this:

$$\begin{aligned} R_\theta^\mathsf{T} R_\theta &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Note: $R_\theta^\mathsf{T} = R_{-\theta}$. This holds for $3D$ rotation matrices also...

Eigenvalues and eigenvectors

If $A \in \mathbb{R}^{n \times n}$ and there is a vector v and scalar λ such that

$$Av = \lambda v$$

Then v is an **eigenvector** of A and λ is the corresponding **eigenvalue**. Some facts:

- Any square matrix has n eigenvalues.
- Each eigenvalue has at least one corresponding eigenvector.
- In general, eigenvalues & eigenvectors can be complex.
- In general, eigenvectors aren't orthogonal, and may not even be linearly independent. i.e. $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ may not be invertible. If it is, we say that A is **diagonalizable** and then $A = V \Lambda V^{-1}$. Otherwise, Jordan Canonical Form.
- Symmetric matrices are **much** simpler!

Recap: symmetric matrices

- Every symmetric $Q = Q^\mathsf{T} \in \mathbb{R}^{n\times n}$ has n real eigenvalues λ_i.
- There exist n mutually orthogonal eigenvectors $u_1, \dots, u_n$:

$$Qu_i = \lambda_i u_i \qquad \text{for all } i = 1, \dots, n$$

$$u_i^\mathsf{T} u_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

- If we define $U = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}$ then $U^\mathsf{T} U = I$ and

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

Eigenvalue example

Consider the quadratic: $7x^2 + 4xy + 6y^2 + 4yz + 5z^2$.
A simple question: are there values that make this negative?

$$\text{equivalent to:} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix}^\mathsf{T} \begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Orthogonal decomposition:

$$\begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}^\mathsf{T}$$

Eigenvalues are $\{3, 6, 9\}$.

Eigenvalue example

Eigenvalue decomposition:

$$\begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}^\mathsf{T}$$

Define new coordinates:

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}^\mathsf{T} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Then we can write:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}^\mathsf{T} \begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}^\mathsf{T} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix}$$

Eigenvalue example

After some manipulations, we discovered that

$$7x^2 + 4xy + 6y^2 + 4yz + 5z^2 \;=\; 3p^2 + 6q^2 + 9r^2$$

where:

$$\begin{aligned} p &= -\tfrac{1}{3}x + \tfrac{2}{3}y - \tfrac{2}{3}z \\ q &= \tfrac{2}{3}x - \tfrac{1}{3}y - \tfrac{2}{3}z \\ r &= \tfrac{2}{3}x + \tfrac{2}{3}y + \tfrac{1}{3}z \end{aligned}$$

Conclusion: the quadratic can never be negative.

Recap

Question: Is $x^\mathsf{T} Q x$ ever negative?

Answer: Look at the orthogonal decomposition of Q:

- $Q = U \Lambda U^\mathsf{T}$
- Define new coordinates $z = U^\mathsf{T} x$.
- $x^\mathsf{T} Q x = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2$

If all $\lambda_i \geq 0$, then $x^\mathsf{T} Q x \geq 0$ for any x.

If some $\lambda_k < 0$, set $z_k = 1$ and all other $z_i = 0$. Then find corresponding x using $x = Uz$, and $x^\mathsf{T} Q x < 0$.

Positive definite matrices

For a matrix $Q = Q^\mathsf{T}$, the following are equivalent:

1. $x^\mathsf{T} Q x \geq 0$ for all $x \in \mathbb{R}^n$

2. all eigenvalues of Q satisfy $\lambda_i \geq 0$

A matrix with this property is called **positive semidefinite** (PSD). The notation is $Q \succeq 0$.

Note: When we talk about PSD matrices, we *always* assume we're talking about a symmetric matrix.

Positive definite matrices

Name	Definition	Notation
Positive semidefinite	all $\lambda_i \geq 0$	$Q \succeq 0$
Positive definite	all $\lambda_i > 0$	$Q \succ 0$
Negative semidefinite	all $\lambda_i \leq 0$	$Q \preceq 0$
Negative definite	all $\lambda_i < 0$	$Q \prec 0$
Indefinite	everything else	(none)

Some properties:

- If $P \succeq 0$ then $-P \preceq 0$
- If $P \succeq 0$ and $\alpha > 0$ then $\alpha P \succeq 0$
- If $P \succeq 0$ and $Q \succeq 0$ then $P + Q \succeq 0$
- Every $R = R^\mathsf{T}$ can be written as $R = P - Q$ for some appropriate choice of matrices $P \succeq 0$ and $Q \succeq 0$.

Ellipsoids

- For linear constraints, the set of x satisfying $c^\mathsf{T} x = b$ is a **hyperplane** and the set $c^\mathsf{T} x \le b$ is a **halfspace**.

- For quadratic constraints:

If $Q \succ 0$, the set $x^\mathsf{T} Q x \le b$ is an **ellipsoid**.

Ellipsoids

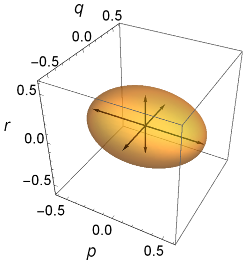
- By orthogonal decomposition, we can write $x^\mathsf{T} Q x = z^\mathsf{T} \Lambda z$ where we defined the new coordinates $z = U^\mathsf{T} x$.
- The set of x satisfying $x^\mathsf{T} Q x \leq 1$ corresponds to the set of z satisfying $\lambda_1 z_1^2 + \cdots + \lambda_n z_n^2 \leq 1$.
- If $Q \succ 0$, then $\lambda_i > 0$. In the z coordinates, this is a stretched sphere (ellipsoid). In the z_i direction, it is stretched by $\frac{1}{\sqrt{\lambda_i}}$.
- Since $x = Uz$, and this transformation preserves angles and distances (think of it as a rotation), then in the x_i coordinates, it is a rotated ellipsoid.
- The principal axes (the z_i directions) map to the u_i directions after the rotation.

Ellipsoids

Plot of the region

$$3p^2 + 6q^2 + 9r^2 \leq 1$$

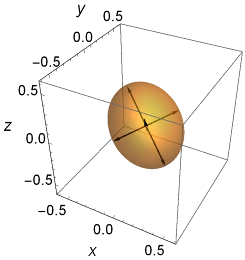
Ellipse axes are in the directions e_1, e_2, e_3



Plot of the region

$$7x^2 + 4xy + 6y^2 + 4yz + 5z^2 \leq 1$$

Ellipse axes are in the directions u_1, u_2, u_3



Norm representation

If $Q \succeq 0$ we can define the **matrix square root**:

1. Let $Q = U\Lambda U^\mathsf{T}$ be an orthogonal decomposition

2. Let $\Lambda^{1/2} = \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$

3. Define $Q^{1/2} = U\Lambda^{1/2}U^\mathsf{T}$.

We have the property that $Q^{1/2}$ is symmetric and $Q^{1/2}Q^{1/2} = Q$. Also:

$$x^\mathsf{T}Qx = (Q^{1/2}x)^\mathsf{T}(Q^{1/2}x) = \left\|Q^{1/2}x\right\|^2$$

Therefore: $x^\mathsf{T}Qx \le b \quad \iff \quad \left\|Q^{1/2}x\right\|^2 \le b$