Lecture Outline

- Convex sets;
- Convex functions;
- Convex optimisation Problems;
- Optimality conditions for convex problems;

You should be able to . . .

- Recognise convex sets and functions
- Identify and formulate convex optimisation problems
- Characterise optimality conditions for convex problems

Affine and Convex Sets

- An *affine set* is a set that contains the line through any two distinct points in it.

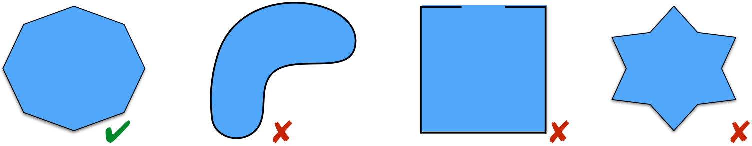
Example: Affine Sets

- A line is an affine set: $x = \theta x_1 + (1 - \theta)x_2, \quad \theta \in \mathbb{R}$
- The solution to $Ax = b$. In fact any affine set can be written in this way.

- A line segment between two points x_1 and x_2 are all points such that: $x = \theta x_1 + (1 - \theta)x_2, \quad \theta \in [0, 1]$.
- A *convex set* contains the line segment between any two points in the set.

$$\forall x_1, x_2 \in \mathcal{S}.\ \theta \in [0, 1] \implies x \in \mathcal{S}, \text{ where } x = \theta x_1 + (1 - \theta)x_2$$

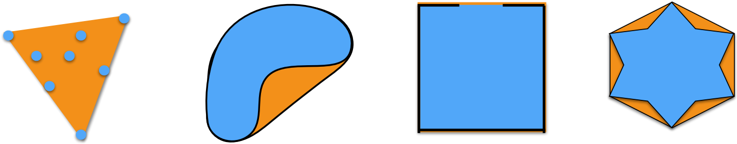
Convex Sets and Convex Hull



- The *convex combination* of n points x_i, $i = 1, \ldots, n$:

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n, \ \theta_1 + \cdots + \theta_n = 1, \ \theta_i \geq 0$$

- The set of all convex combinations of members of a set is the *convex hull* of the set.



Conic Combination and Cones

- *Conic combination* of two points x_1 and x_1:

$$x = \theta_1 x_1 + \theta_2 x_2, \quad \theta_1, \theta_2 \geq 0$$

- A *convex cone* is a set that contains all conic combinations of the points in the set.
- A convex cone is a *proper cone* if
  1. it is closed (contains its boundary);
  2. it is solid (has nonempty interior);
  3. it is pointed (contains no lines).

Example: Proper Cones

- Nonnegative orthant: $\{x | x_i \geq 0, i = 1, \ldots, n\}$
- Positive semidefinite cone: $\{X \in \mathbb{R}^{n \times n} | X = X^T,\ v^T X v \geq 0, \forall v \neq 0\}$

Excursion: (Semi)Definite Matrices

- A symmetric matrix P, $P = P^T$, is positive (semi-)definite, $P \succ 0$ ($P \succeq 0$) if and only if
  - $\forall v \neq 0$, $v^T P v > 0$ ($v^T P v \geq 0$),
  - or equivalently $\lambda_{\min}(P) > 0$ ($\lambda_{\min}(P) \geq 0$)
- Generalised inequality for symmetric matrices:

$$P \succeq Q \iff P - Q \succeq 0$$

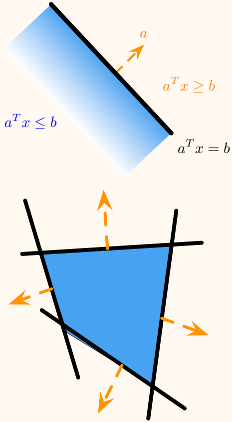
- For the ease of presentation $\succ$, $\succeq$, $\prec$, and $\preceq$ are replaced by $>$, $\geq$, $<$, and $\leq$ when the operands are symmetric matrices.
- For vectors x and y we define the generalised inequality with respect to the positive orthant:

$$x \leq y \iff x_i \leq y_i, \quad i = 1, \ldots, n$$

Convex Sets

Example: Convex Sets

- Hyperplane: $\{x | a^T x = b\}$ (affine)
- Halfspace: $\{x | a^T x \leq b\}$
- Polyhedrals/Polygons: intersection of halfspaces and hyperplanes, $\{x | Ax \leq b, Cx = d\}$
- Norm balls with centre x_c and radius r: $\mathcal{B}(x_c, r) = \{x | \|x - x_c\| \leq r\}$.
- Norm cone: $\{(x, t) | \|x\| \leq t\}$. If the Euclidean norm is used the cone is called the second order cone.



Convex Sets

Example: Convex Sets

- Ellipsoid: $\{x|(x-x_c)^T P^{-1}(x-x_c) \leq 1\}$, where $P > 0$.
- Alternative representation of an ellipsoid ($A = P^{1/2}$):

$$\{x_c + Au | \|u\| \leq 1\}$$

- The intersection of finitely or infinitely many convex sets.
- The affine image of a convex set is convex:

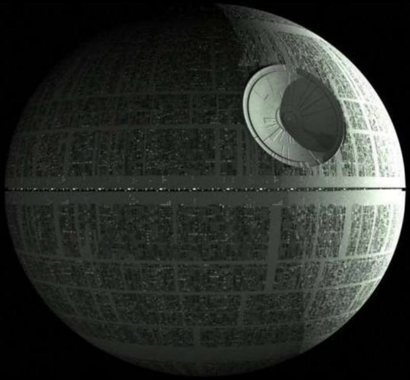
$$A\mathcal{S} + b = \{y|\exists x \in \mathcal{S}, y = Ax + b\}$$

- The affine pre-image of a convex set is convex:

$$\{y|Ay + b \in \mathcal{S}\}$$

- The sublevel sets of a convex function are convex

A moon: convex



Not a moon: nonconvex

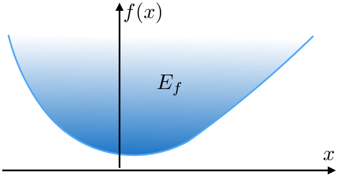
Convex Functions

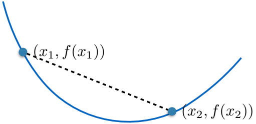
- A function $f : \mathcal{D} \to \mathbb{R}$ is *convex* if $\mathcal{D}$ is convex and

$$\forall x, y \in \mathcal{D}, \theta \in [0,1] : f((1-\theta)x + \theta y) \leq (1-\theta)f(x) + \theta f(y)$$

- Define the *epigraph* of f as

$$\text{epi}(f) = \{(x,s) | x \in \mathcal{D}, s \geq f(x)\}$$





- If strict inequality holds the function is *strictly convex*
- All secants of a convex function are above the graph.
- The epigraph of a function is convex iff the function is convex.

Convex Functions

- The notion of *strong convexity* extends and parametrises strict convexity. Any strongly convex function is strictly convex, not vice versa.

Definition (Strong Convexity): *A function* $f : \mathcal{D} \to \mathbb{R}$ *is strongly convex with coefficient (modulus)* σ *if* $\mathcal{D}$ *is convex and*

$$f((1-\theta)x + \theta y) + \frac{\sigma}{2}\theta(1-\theta)\|x-y\|^2 \leq (1-\theta)f(x) + \theta f(y)$$

for all $\theta \in [0,1]$ *and* $x, y \in \mathcal{D}$, *and some* $\sigma > 0$.

- Also if f is differentiable:
- $(\nabla f(x) - \nabla f(y))^T (x-y) \geq \sigma \|x-y\|^2$
- $f(y) \geq f(x) + \nabla f(x)^T (y-x) + \frac{\sigma}{2}\|x-y\|^2$
- $\nabla^2 f(x) - \sigma I \geq 0$

Convex Optimisation

Definition (Convex Optimisation Problem): *The optimisation problem*

$$\min \quad f(x), \quad s.t. \quad x \in \mathcal{X}$$

is called a convex optimisation problem, if f is a convex function and $\mathcal{X}$ is a convex set.

- A function f is *concave* if $-f$ is convex.
- The minimisation problem above is equivalent to

$$\max \quad -f(x), \quad \text{s.t.} \quad x \in \mathcal{X}$$

Theorem (Local Implies Global Optimality for Convex Problems): *Consider the convex optimisation problem described above. Then every local minimum is also a global one.*

Convex Optimisation

- Let $x^\star$ be a local minimum. We show for any $y \in \mathcal{X}$, $f(x^\star) \leq f(y)$.
- Consider the neighbourhood $\mathcal{N}$ of $x^\star$ such that $f(x^\star) \leq f(\bar{x}), \forall \bar{x} \in \mathcal{X} \cap \mathcal{N}$.
- Note that the line segment between $x^\star$ and y lies in $\mathcal{X}$.
- Consider an arbitrary $z \in \mathcal{N}$ where $z = (1-\theta)x^\star + \theta y$ for some $\theta \in [0,1]$.

$$f(x^\star) \leq f(z) = f((1-\theta)x^\star + \theta y) \leq (1-\theta)f(x^\star) + \theta f(y) \implies 0 \leq \theta(f(y) - f(x^\star)) \implies f(x^\star) \leq f(y). \qquad \square$$

- If the function is strictly convex in the neighbourhood of the solution, the minimiser is unique (strict inequality above).
- For a convex problem either there is a unique minimiser or the minimisers form a convex set.

Convex Optimisation

Theorem (Convexity of the Set of Minimisers): *Consider the convex optimisation problem described above. Then the set of its minimisers is convex.*

- Assume the minimum $f^\star$ is attained at points x_1 and x_2.
- For any $z = (1-\theta)x_1 + \theta x_2$, $\theta \in [0,1]$:

$$f^\star \leq f(z) = f((1-\theta)x_1 + \theta x_2) \leq (1-\theta)f^\star + \theta f^\star \leq f^\star$$

- Thus, z is a minimiser as well and the set of all minimisers is convex.

Convex Optimisation

Theorem (Convexity for Differentiable Functions): *Assume a continuously differentiable function $f : \mathcal{D} \to \mathbb{R}$ with convex domain. It is convex if and only if:*

$$\forall x, y \in \mathcal{D} : \quad f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

In other words, tangents lie below the graph.

- "If convex":

$$\begin{aligned} f((1-\theta)x + \theta y) =& f(x + \theta(y-x)) \leq f(x) + \theta(f(y) - f(x)) \\ & f(x + \theta(y-x)) - f(x) \leq \theta(f(y) - f(x)) \end{aligned}$$

$$\nabla f(x)^T (y - x) = \lim_{\theta \to 0} \frac{f(x + \theta(y-x)) - f(x)}{\theta} \leq f(y) - f(x)$$

- "if the inequality": $z = (1-\theta)x + \theta y$

$$f(z) + \nabla f(z)^T (x - z) \leq f(x), \quad f(z) + \nabla f(z)^T (y - z) \leq f(y)$$

$$f(z) + \nabla f(z)^T \underbrace{\left[(1-\theta)(x-z) + \theta(y-z)\right]}_{=0} \leq (1-\theta) f(x) + \theta f(y)$$

Convex Optimisation

Theorem (Convexity for Twice Differentiable Functions): *Assume a twice continuously differentiable function $f : \mathcal{D} \to \mathbb{R}$ with convex domain. It is convex if and only if:* $\forall x \in \mathcal{D}: \quad \nabla^2 f(x) \geq 0.$

- "If convex": From Taylor's expansion

$$f(x+tp) = f(x) + t\nabla f(x)^T p + \frac{1}{2} t^2 p^T \nabla^2 f(x) p + o(t^2 \|p\|^2) \implies$$

$$p^T \nabla^2 f(x) p = \lim_{t \to 0} \frac{2}{t^2} \underbrace{\left[f(x+tp) - f(x) - t\nabla f(x)^T p\right]}_{\geq 0, \text{from the previous Thm}} \geq 0$$

- "if the inequality": Taylor's theorem, $p = (y-x)$, $t \in [0,1]$

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2} \underbrace{(y-x)^T \nabla^2 f(x+t(y-x))(y-x)}_{\geq 0}$$

$$\implies f(y) \geq f(x) + \nabla f(x)^T (y-x)$$

Convex Functions

Example: Convex Functions

- Exponential function: $f(x) = e^x$
- Quadratic function: $f(x) = c^T x + \frac{1}{2} x^T Q x, \quad Q \geq 0$
- $f(x,t) = \frac{x^T x}{t}$, $\mathcal{D} = \{(x,t) | x \in \mathbb{R}^n, t > 0\}$
- Affine function: $f(x) = a^T x + b$
- p-norm: $\|x\|_p = \left(\sum_{i=1}^p |x_i|^p\right)^{1/p}$, $p \geq 1$
- ∞-norm: $\|x\|_\infty = \max_{k \in \{1,\dots,n\}} |x_k|$
- $f(X) = \text{trace}(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$, $X \in \mathbb{R}^{m \times n}$
- $f(X) = \|X\|_2 = \sigma_{\max}(X) = \sqrt{\lambda_{\max}(X^T X)}$, $X \in \mathbb{R}^{m \times n}$

Convex Functions

Theorem: *The function $f : \mathcal{D} \to \mathbb{R}$ is convex if and only if $g : \mathcal{D}_g \to \mathbb{R}$ where $g(t) = f(x+tp)$ and $\mathcal{D}_g = \{t | x+tp \in \mathcal{D}\}$ is convex for any line restriction, i.e. for any $x \in \mathcal{D}$ and $p \in \mathbb{R}^n$.*

- This can be used to show the convexity of $f(X) = -\log\det(X),\ X > 0$:

$$g(t) = -\log\det(X + tP) = -\log\det(X) - \log\det(I + tY)$$
$$= -\log\det(X) - \sum_{i=1}^{n} \log(1 + t\lambda_i)$$

- And $g(t)$ is convex where $Y = X^{-1/2} P X^{-1/2}$ and $\lambda_i = \lambda_i(Y)$.

Operations That Preserve Convexity

- Input Transformation: If $f : \mathcal{D} \to \mathbb{R}$ is convex then $g(x) = f(Ax + b)$ is also convex on $\mathcal{D}_g = \{x | Ax + b \in \mathcal{D}\}$.
- Extension: If $f : \mathcal{D} \to \mathbb{R}$ is convex, then $g(x) = \begin{cases} f(x) & x \in \mathcal{D} \\ +\infty & \text{otherwise} \end{cases}$ is convex.
- Summation: Summation of convex functions or k largest entries of a vector.
- Point-wise Maximum: If $f_1, \ldots, f_m$ are convex then $f(x) = \max_{i \in \{1,\ldots,m\}} \{f_1(x), \ldots, f_m(x)\}$ is convex.
- Point-wise Supremum: If $f(x, y)$ is convex in x for $y \in \mathcal{Y}$, then

$$g(x) = \sup_{y \in \mathcal{Y}} f(x, y)$$

is convex. For example,

  - Distance to the farthest point in a set: $\sup_{y \in \mathcal{Y}} \|x - y\|$
  - The maximum eigenvalue of a symmetric matrix: $\lambda_{\max}(X) = \sup_{\|y\|=1} y^T X y$

Operations That Preserve Convexity

- Minimisation: If $f(x, y)$ is convex in (x, y) and $\mathcal{S}$ is convex then so is

$$g(x) = \min_{y \in \mathcal{S}} f(x, y)$$

  - Distance to a convex set: $\min_{y \in \mathcal{Y}} \|x - y\|$
- Composition: The function $f = h(g(x))$ is convex if
  - g is convex, h is convex and nondecreasing
  - g is concave, h is convex and nonincreasing
  - These conditions can be observed from:

$$\nabla^2 f(x) = h''(g(x)) \nabla g(x) \nabla g(x)^T + h'(g(x)) \nabla^2 g(x) \geq 0$$

- Perspective: If $f : \mathcal{D} \to \mathbb{R}$ is convex then $g(x, t) = t f(x/t)$ is convex over $\{(x, t) | x/t \in \mathcal{D}, t > 0\}$.

Standard Form of Convex Optimisation Problems

$$
\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & c_i(x) = 0, \quad i \in \mathcal{E} \\
& c_i(x) \geq 0, \quad i \in \mathcal{I}
\end{aligned}
$$

Sufficient condition for it to be convex:

- f is convex.
- $c_i(x)$ is affine for $i \in \mathcal{E}$.
- $c_i(x)$ is concave $i \in \mathcal{I}$.
- The affine constraints are often just replaced by $Ax = b$.

$$
\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & Ax = b \\
& c_i(x) \geq 0, \quad i \in \{1, \ldots, m\}
\end{aligned}
$$

Famous Convex Optimisation Problems

Example: Quadratically Constrained Quadratic Program (QCQP) and Semidefinite Program (SDP)

$$\text{QCQP:} \quad \min \quad \frac{1}{2}x^T Q_0 x + q_0^T x$$
$$\text{s.t.} \quad Ax = b$$
$$\frac{1}{2}x^T Q_i x + q_i^T x + d_i \leq 0, \quad i = 1, \dots, m$$

where $Q_i = Q_i^T \geq 0$. Linear Programmes (LP) and Quadratic Programmes (QP) are special cases.

$$\text{SDP:} \quad \min \quad q^T x$$
$$\text{s.t.} \quad Ax = b$$
$$P_0 + \sum_{i=1}^{m} x_i P_i \geq 0, \text{ (linear matrix inequality (LMI))}$$

where $P_i^T = P_i \geq 0$. QCQPs can be written as SDPs.

Famous Convex Optimisation Problems

Example: Quadratically Constrained Quadratic Program (QCQP) as Semidefinite Program (SDP)

$$
\begin{aligned}
\min \quad & t \\
\text{s.t.} \quad & Ax = b \\
& \frac{1}{2}x^T Q_0 x + q_0^T x - t \leq 0, \\
& \frac{1}{2}x^T Q_i x + q_i^T x + d_i \leq 0, \quad i = 1, \ldots, m
\end{aligned}
$$

- Note that $Q_i = R_i^T R_i$. From the Schur complement:

$$
\frac{1}{2}x^T Q_i x + q_i^T x + d_i \leq 0 \iff \begin{bmatrix} -I & R_i x \\ x^T R_i^T & q_i^T x + d \end{bmatrix} \leq 0
$$

$$
\frac{1}{2}x^T Q_0 x + q_0^T x - t \leq 0 \iff \begin{bmatrix} -I & R_0 x \\ x^T R_0^T & q_0^T x - t \end{bmatrix} \leq 0
$$

- Thus the constraints can be written as LMI.

Excursion: Schur Complement Conditions

Schur complement conditions for positive definiteness

- $X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$
- $X > 0 \iff A > 0,\ C - B^T A^{-1} B > 0$
- $X > 0 \iff C > 0,\ A - B C^{-1} B^T > 0$
- If $A > 0$ then $X \geq 0 \iff A > 0,\ C - B^T A^{-1} B \geq 0$
- If $C > 0$ then $X \geq 0 \iff C > 0,\ A - B C^{-1} B^T \geq 0$

- Norm Constraints as SDP (X a vector or a matrix):

$$\|X\| \leq t \iff \begin{bmatrix} tI & X \\ X^T & tI \end{bmatrix} \geq 0$$

- Eigenvalue Constraints as SDP:

$$\lambda_{\max}(X) \leq t \iff tI - X \geq 0$$

Convex Optimisation Problems

Theorem (First Order Optimality Condition for Convex Problems): *Consider the convex optimisation problem with continuously differentiable objective function*

$$\min \quad f(x), \quad s.t. \quad x \in \mathcal{X}.$$

A point $x^\star$ is a global minimiser if and only if

$$\forall x \in \mathcal{X}, \ \nabla f(x^\star)^T (x - x^\star) \geq 0.$$

- "$\Rightarrow$": To obtain a contradiction assume $\exists x$, $\nabla f(x^\star)^T (x - x^\star) < 0$. Let $z = x^\star + \theta(x - x^\star)$. From the Taylor expansion:

$$f(z) = f(x^\star) + \theta \nabla f(x^\star)^T (x - x^\star) + o(\theta) < f(x^\star)$$

- Which is a contradiction for small $\theta > 0$.
- "$\Leftarrow$": $f(x) \geq f(x^\star) + \nabla f(x^\star)^T (x - x^\star) \geq f(x^\star)$

Convex Optimisation Problems

Theorem (Unconstrained Convex Problems): *Consider the convex optimisation problem with a twice continuously differentiable objective function,* $\min \quad f(x)$*. A point* $x^\star$ *is a global optimiser if and only if* $\nabla f(x^\star) = 0$*.*

Example: Strictly Convex Unconstrained Quadratic

- Consider $(Q > 0)$

$$\min \quad f(x) = \frac{1}{2} x^T Q x + q^T x$$

- Applying the theorem:

$$\nabla f(x) = Qx + q = 0 \implies x^\star = -Q^{-1} q$$

$$f(x^\star) = -\frac{1}{2} q^T Q^{-1} q$$