Session - 7, Feb 19, 2025

Chapter - 4

Convex optimization problem

standard form convex optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & \underline{a_i^T x = b_i,} \quad i = 1, \dots, p \end{array}$$

- $f_0, f_1, \dots, f_m$ are convex; equality constraints are affine
- problem is *quasiconvex* if f_0 is quasiconvex (and $f_1, \dots, f_m$ convex)

often written as

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b \end{array}$$

important property: feasible set of a convex optimization problem is convex

example

Max e^x, $x \in [0,1]$

$$\begin{array}{ll} \text{minimize} & f_0(x) = x_1^2 + x_2^2 \\ \text{subject to} & f_1(x) = x_1/(1+x_2^2) \leq 0 \\ & h_1(x) = (x_1+x_2)^2 = 0 \end{array}$$

check that it is a feasible set.

- f_0 is convex; feasible set $\{(x_1,x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (according to our definition): f_1 is not convex, h_1 is not affine

~~check it~~ check it

- equivalent (but not identical) to the convex problem

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \leq 0 \\ & x_1 + x_2 = 0 \end{array}$$

Problem: minimize $f_0(x_1, x_2)$

subject to
$$\left|\begin{array}{l} 2x_1 + x_2 \geq 1 \\ x_1 + 3x_2 \geq 1 \\ x_1 \geq 0,\ x_2 \geq 0 \end{array}\right.$$

Make a sketch of feasible set.

Find all (x_1, x_2) which satisfies the constraints.

Step 1: Identify the constraints.

St-2 $\rightarrow$ Convert the inequalities to equality.

St. 3 $\Rightarrow$ Find intersection of two lines $2x_1 + x_2 = 1$, $x_1 + 3x_2 = 1$.

Intersection point is $\left(\frac{2}{5}, \frac{1}{5}\right)$ ✓

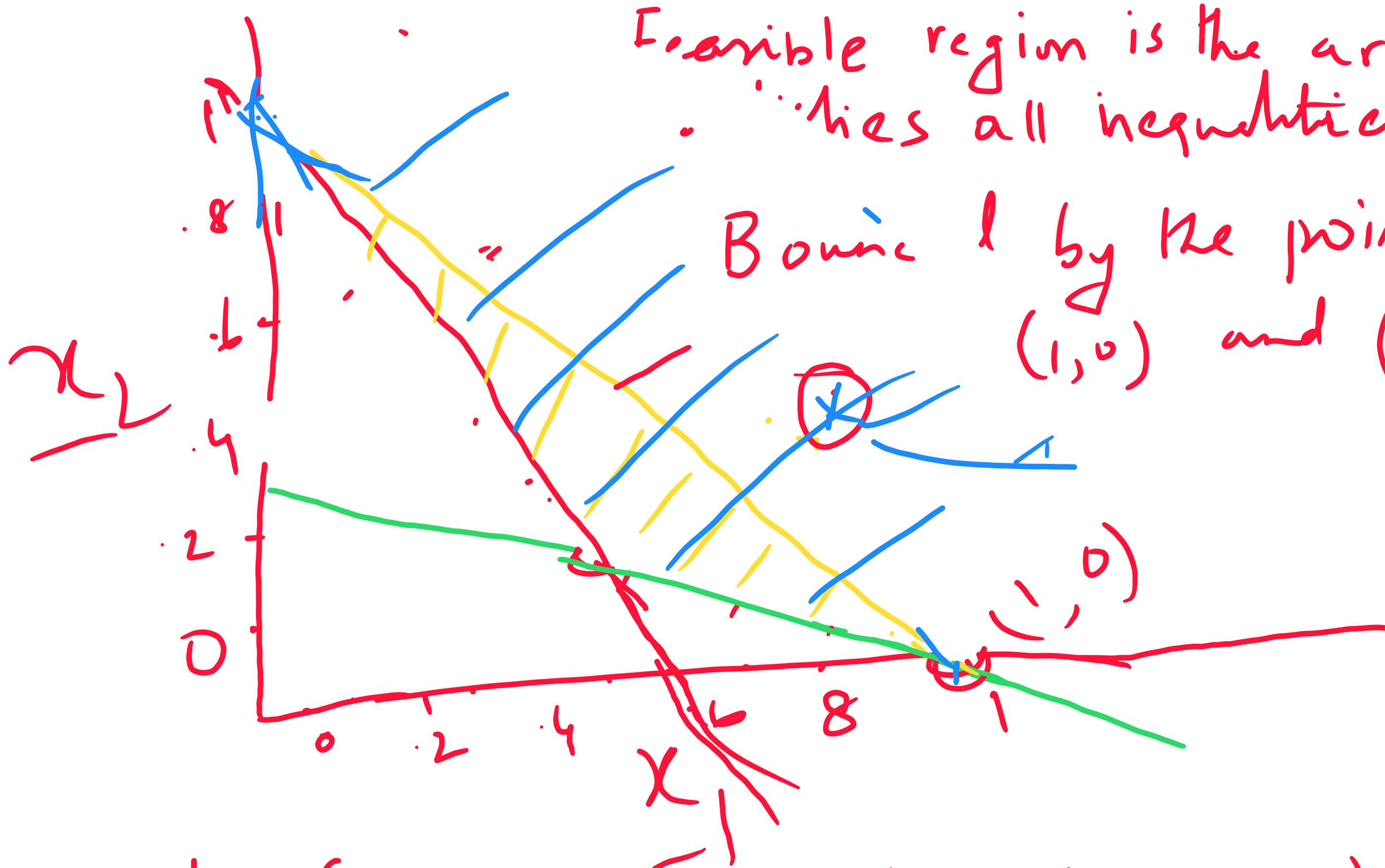
S-4: x-y axis intersection. $x_1 = 0$ in $2x_1 + x_2 = 1 \Rightarrow x_2 = 1$.

$x_2 = 0$ in $x_1 + 3x_2 = 1 \Rightarrow x_1 = 1$

Points are $(1,0)$ and $(0,1)$

Feasible region is the area that satisfies all inequalities.

Bounded by the points $(0,1)$, $(1,0)$ and $\left(\frac{2}{5}, \frac{1}{5}\right)$



Step 4: Compute $f_0(x_1, x_2)$ at $(0,1)$, $(1,0)$, $(2/5, 1/5)$.

H.W: Minimise $f_0(x_1, x_2) = -x_1 - x_2$

Minimise $f_0(x_1, x_2) = \max(x_1, x_2)$

Ref: 4.2.2, page 138

Local and global optima

any locally optimal point of a convex problem is (globally) optimal

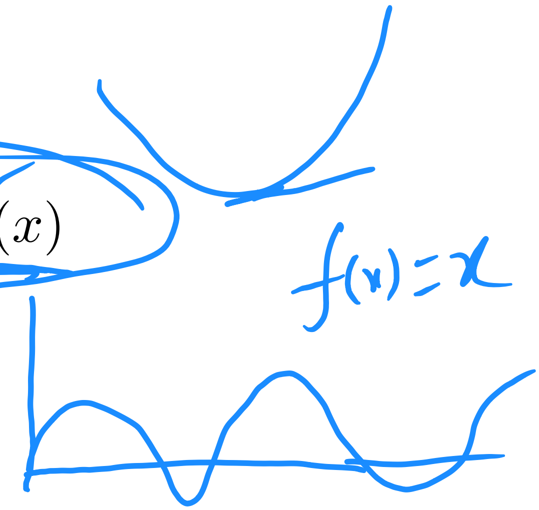
proof: suppose x is locally optimal, but there exists a feasible y with

$f_0(y) < f_0(x)$ x is not global optima.

x locally optimal means there is an $R > 0$ such that

given

$$z \text{ feasible}, \quad \|z - x\|_2 \leq R \quad \Longrightarrow \quad f_0(z) \geq f_0(x)$$



$f(x) = x$

Given

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and $\|\theta y + (1-\theta)x - x\| = \theta\|y - x\|_2$

$$f_0(z) \leq \theta f_0(y) + (1 - \theta)f_0(x) < f_0(x)$$

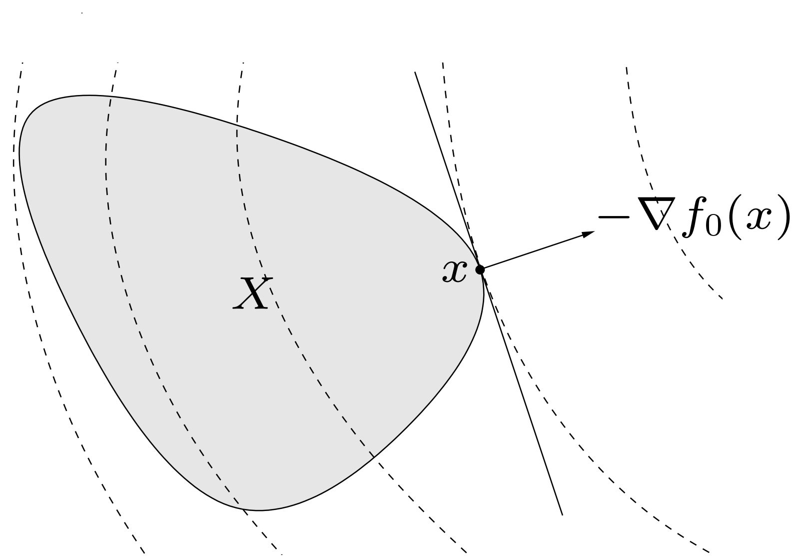
$f_0(x) + \theta(f_0(y) - f_0(x))$

which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T(y - x) \geq 0 \quad \text{for all feasible } y \qquad (1)$$



if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

Prove that $x^* = (1, 1/2, -1)$ is optimal for the optimization problem

$$\text{minimize } (1/2)x^T P x + q^T x + r$$

p.d

subject to $-1 \le x_i \le 1, \ i = 1, 2, 3.$

where
$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \quad q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13.0 \end{bmatrix}, \quad r = 1$$

Verify that x^* satisfies the above optimality condition as discussed in the class.

Due: Gradient derivation (next class)

Session End

Session - 8

Gradient

Let f be a real valued function ($f : \mathbb{R}^n \to \mathbb{R}$)

the dirivative of f.

Df is a $\underline{(1 \times n)}$ matrix.
row matrix.

$f(x_1, x_2) = x_1^2 + x_2^2$,

$$Df = (2x_1, 2x_2)$$

Gradient of f is the transpose of Df

$$\nabla f(x) = Df(x)^T$$

It is a column vector $(n \times 1)$

Components of $\nabla f(x)$ would be

$$(\nabla f(x))_i = \frac{\partial f(x)}{\partial x_i}, \quad i = 1(1)n.$$

Example: $f(x) = (1/2) x^T P x + q^T x + r_0$

$P \in S^n, q \in \mathbb{R}^n, r \in \mathbb{R}.$

Compute gradient of $f(x)$.

Derivative at 'x'

Row vector $Df(x) = x^T P^{n \times n} + q^T$

Column vector $\nabla f(x) = Df(x)^T$

$= Px + \underline{q}$

Unconstrained problem:

$m = p = 0$

$f_0(\underline{x})$ is convex, optimize it

If f_0 is differentiable,

then x would be optimal

if $\nabla f_0(x) = 0$ — (2)

Show that (2) can be derived from

(1).

Ref: 4.2.3, page 140.

- **unconstrained problem**: x is optimal if and only if

$$x \in \mathbf{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

- **equality constrained problem**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad Ax = b$$

x is optimal if and only if there exists a ν such that

$$x \in \mathbf{dom}\, f_0, \qquad Ax = b, \qquad \boxed{\nabla f_0(x) + A^T \nu = 0}$$

- **minimization over nonnegative orthant**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad x \succeq 0$$

(x_1, x_2, x_3)

x is optimal if and only if

$$x \in \mathbf{dom}\, f_0, \qquad x \succeq 0, \qquad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

Equivalent convex problems

two problems are (informally) **equivalent** if the solution of one is readily obtained from the solution of the other, and vice-versa

some common transformations that preserve convexity:

- **eliminating equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & \underline{Ax = b} \end{array}$$

  is equivalent to

$$\begin{array}{ll} \text{minimize (over } z\text{)} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{array}$$

  where F and x_0 are such that

$$\boxed{Ax = b \quad \Longleftrightarrow \quad x = Fz + x_0 \text{ for some } z}$$

- **introducing equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(\overset{y}{A_0x} + b_0) \\ \text{subject to} & f_i(A_ix + b_i) \leq 0, \quad i = 1, \dots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x, y_i) & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, \quad i = 1, \dots, m \\ & y_i = A_ix + b_i, \quad i = 0, 1, \dots, m \end{array}$$

- **introducing slack variables for linear inequalities**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^Tx \leq b_i, \quad i = 1, \dots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x, s) & f_0(x) \\ \text{subject to} & a_i^Tx + s_i = b_i, \quad i = 1, \dots, m \\ & s_i \geq 0, \quad i = 1, \dots m \end{array}$$

- **epigraph form**: standard form convex problem is equivalent to

$$\begin{array}{ll} \text{minimize (over } x, t) & t \\ \text{subject to} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- **minimizing over some variables**

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x_1) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

show that the following are equivalent problems:

Suppose A (Data set) $\in R^{m \times n}$, a_i^T is row.

vector $b \in \mathbb{R}^m$, M constant >0

(a) Robust least square problem.

$$\text{Minimize } \sum_{i=1}^{m} \phi(a_i^T x - b_i)$$

with $x \in \mathbb{R}^n$, and $\phi: R \to R$

(Huber penalty function)

$$\phi(u) = \begin{cases} u^2 & \text{if } |u| \leq M \\ M(\underline{2|u| - M}) & |u| > M \end{cases}$$

(b) LS with weight

$$\text{minimize } \sum_{i=1}^{m} (a_i^T x - b_i)^2 / (w_i + 1)$$

$$+ \frac{\mu^2}{2} \mathbf{1}^T w$$

subject to $w \geq 0$.

Ref: Exercise 4.5 of the textbook, page 190.