Lecture 23

A general form of optimization problems.

① Constrained optimization:

$$\text{Min (or max) } f(x) \qquad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$
$$\text{s.t.} \quad g_i(x) \geq 0 \quad i = 1, \dots, p$$
$$h_j(x) = 0 \quad j = 1, \dots, q.$$

Where f, g, h are <u>any</u> functions of x.

- In linear programming, $f(x), g(x), h(x)$ are <u>linear functions</u> of x.

<u>Ex</u>: $f(x) = c_1x_1 + \dots + c_nx_n$
$g_i(x) = x_i$

and

$h(x) = Ax - b$.

$\Downarrow$

$$\text{Max } c^t x$$
$$\text{s.t.} \quad x_1 \geq 0$$
$$\vdots$$
$$x_n \geq 0.$$
$$Ax - b = 0$$

$x = (x_1, x_2)$

Ex: $f(x) = 3x_1 + 5x_2$.
$g_3(x) = x_1$
$g_4(x) = x_2$.
$g_1(x) = 5x_1 + 7x_2 - 20$
$g_2(x) = -4x_1 + 2x_2 - 17$
$h_1(x) = 3x_1 - 6x_2 = 10$

$\Downarrow$

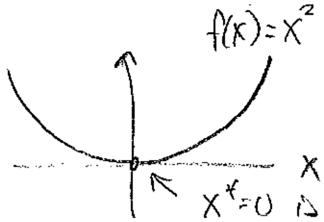
$$\text{Min } 3x_1 + 5x_2$$
$$\text{s.t.} \quad 5x_1 + 7x_2 \geq 20$$
$$-4x_1 + 2x_2 \geq 17$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$
$$3x_1 - 6x_2 = 10.$$

② Unconstrained optimization:

$$\underset{x \in \mathbb{R}^n}{\text{Min (or max)}} \ f(x)$$

where f is any function of x.

<u>Ex</u>: $\underset{x \in \mathbb{R}^1}{\text{Min}} \ x^2$



<u>Remark</u>: • When the functions f, g, h are "nice" (e.g. linear), the problem is easier to solve.

In particular, when f, g, h are linear, we have "linear programming" which we have studied extensively in Opt 1.

Our goals this week is to consider nonlinear programming problems of the following standard form:

$$\left.\begin{array}{lll} \text{Min} & f(x) & \\ \text{s.t.} & g_i(x) \geq 0 & i = 1, \dots, p \\ & h_j(x) = 0 & j = 1, \dots, q. \end{array}\right\} (NLP)$$

where $x = (x_1, \dots, x_n)$

$f(x)$, $g_i(x)$, $h_j(x)$ are <u>continuous</u> and <u>differentiable</u> functions of $x_1, \dots, x_n$.

<u>Ex</u>:

$$f(x_1) = \log x_1 \qquad n = 1.$$
$$g(x_1) = x_1 - 2$$

<u>Ex</u> : (not continuous).

$$f(x_1) = \mathbb{1}\{x_1 \geq 3\}$$

(continuous but not differentiable)

$$f(x_1) = |x_1 - 3|.$$

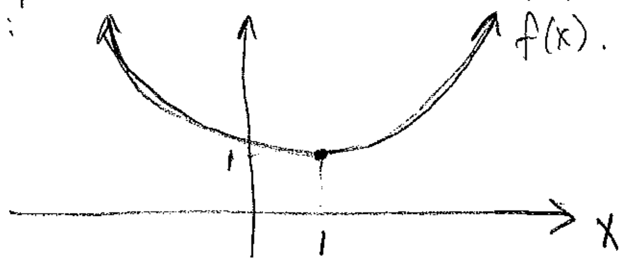
Let's try to gain some intuition for obtaining a method for solving nonlinear optimization problems:

I. Unconstrained Optimization Consider an example of an unconstrained optimization problem with nonlinear objective function:

Ex 1: $\min_{x \in \mathbb{R}^1} \quad x^2 - 2x + 2$.

- Here, $f(x) = x^2 - 2x + 2$ is nonlinear in x.

Plot:



1-clicker

(From this picture, $x=1$ seems to be the optimal solution, with value $f(1) = 1$.)

- A more methodical approach from calculus: Differentiate f wrt x:

$$f'(x) = 2x - 2$$

and solve for x s.t. $f'(x) = 0$:

$$0 = 2x - 2.$$
$$\therefore \; 2 = 2x$$
$$\therefore \; x = 1$$

→ At x, $f'(x)$ is the slope of f.
If $f'(1) = 0$,
then $x=1$ is a point where the value of f does not increase or decrease with small perturbation around x.

So, we propose the following condition:

Proposition? Given a function $f(x)$ (in one variable) that is continuous and differentiable, if $f'(\bar{x}) = 0$, then $\bar{x}$ is a minimizer (an opt soln) of f.

Ex 2: $\underset{x \in \mathbb{R}^1}{\text{Minimize}} \quad 35 - 12x^2 + 4x^3 + 3x^4$

Here, $f(x) = 35 - 12x^2 + 4x^3 + 3x^4$

- Differentiating f:

$$\begin{aligned} f'(x) &= -24x + 12x^2 + 12x^3 \\ &= 12x(-2 + x + x^2) \\ &= 12x(x+2)(x-1). \end{aligned}$$

So, $f'(x) = 0$ when $x = 0$, $x = -2$, or $x = 1$.

Are they all minimizers of f?
Check:

$$\begin{aligned} f(0) &= 35 \\ f(-2) &= 35 - 48 - 32 + 48 = 3. \leftarrow \text{the smallest} \\ f(1) &= 35 - 12 + 4 + 3 = 30 \end{aligned}$$

No! So, our proposition is not correct!

In fact, from calculus,
if x satisfies $f'(x) = 0$,
① then x is either a local minimizer, a local maximizer, or an inflection point.
② x is a local minimizer only if $f''(x) > 0$.

- The second derivative of f:

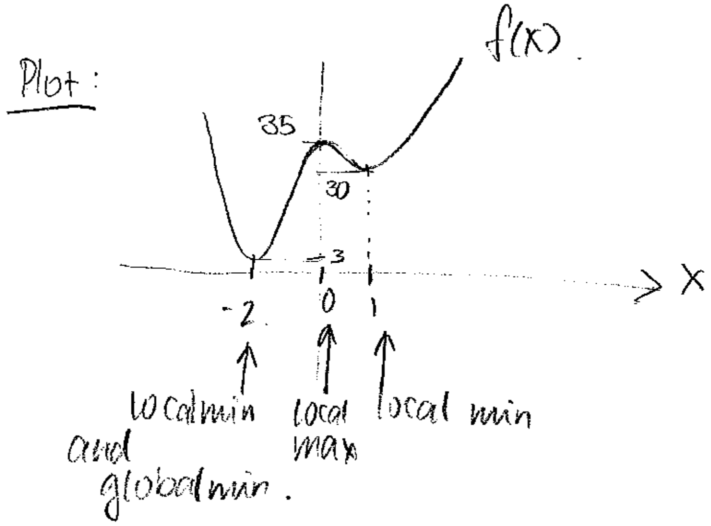
$$f''(x) = -24 + 24x + 36x^2$$

So, $f''(0) = -24 < 0 \rightarrow$ (local maximizer)

$f''(1) = -24 + 24 + 36 = 36 > 0$.
$f''(-2) = -24 - 48 + 144 = 72 > 0$

$\therefore$ $x = 1$ and $x = -2$ are both local minimizers!

Plot:



So, what have we observed so far?

For the problem $\min_{x \in \mathbb{R}^n} f(x)$ } (NLP).

Observations: For a continuous and differentiable function $f(x)$ (single-variable):

① A point $\bar{x}$ is a local minimizer of f if:

$$f'(\bar{x}) = 0$$
$$f''(\bar{x}) > 0$$

[A point $\bar{x}$ is a local maximizer of f if:

$$f'(\bar{x}) = 0$$
$$f''(\bar{x}) < 0.]$$

Thm: ② If $\bar{x}$ is a (global) minimizer of f, then it satisfies:

$$\left.\begin{aligned} f'(\bar{x}) &= 0 \\ f''(\bar{x}) &> 0, \end{aligned}\right\}$$ necessary conditions for a minimizer of f.

but if $\tilde{x}$ satisfies $f'(\tilde{x}) = 0$, $f''(\tilde{x}) > 0$, it is not necessarily a (global) maximizer of f.

However, for a special type of continuous & differentiable function f, we can more easily find the (global) minimizer:

Thm If f is a continuous, differentiable, and convex function (in one variable) then: if $\bar{x}$ satisfies $f'(\bar{x}) = 0$, then $\bar{x}$ is a (global) minimizer of f.

<u>Defn</u> A function f is <u>convex</u> if:

given 2 points a and b where $a<b$,
and any number $0 \le \lambda \le 1$, then

$$f(\underbrace{\lambda a + (1-\lambda)b}_{\text{a point between } a \text{ and } b}) \le \lambda f(a) + (1-\lambda)b.$$

<u>Ex</u>: $f(x) = x^2 - 2x + 2$. is convex.

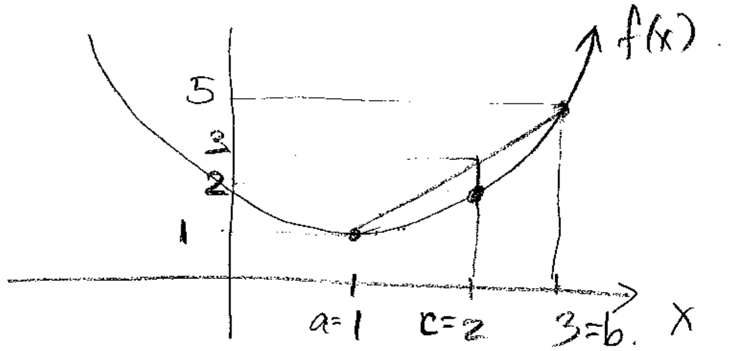
Let $a=1$, $b=3$ (for example).
$\lambda = \frac{1}{2}$;

Let $c := \frac{1}{2}a + (1-\frac{1}{2})b = \frac{1}{2} + \frac{1}{2} \cdot 3 = 2.$,

$f(c) = f(2.) = 4 - 4 + 2 = 2.$

$\frac{1}{2}f(a) + (1-\frac{1}{2})f(b) = \frac{1}{2}(1) + \frac{1}{2}(5) = 3.$

$\therefore f(\lambda a + (1-\lambda)b) = 2 \le 3 = \lambda f(a) + (1-\lambda)f(b).$



i.e. f is convex if $(c, f(c))$ is under the line connecting $(a, f(a))$ and $(b, f(b))$

(*) → Summary (boxd)

<u>Remark</u>: • If f is <u>continuous</u>, <u>differentiable</u>, and <u>convex</u>, then $f''(x) \ge 0$ for all x.

• If f has a local minimizer, it is <u>the only one</u>, so it is also a global minimizer.

→ Summary (boxed) → necessary and optimal conditions

ASK Q: Our theorem also holds for convex functions of several variables.

Ex: Minimize $2x_1^2 - 2x_1x_2 + 3x_2^2 + 2x_1 + 1$.

$$f(x) = f(x_1, x_2) = 2x_1^2 - 2x_1x_2 + 3x_2^2 + 2x_1 + 1.$$

(Claim: $f(x)$ is convex → see plot)

The gradient of f:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 4x_1 - 2x_2 + 2 \\ -2x_1 + 6x_2 \end{pmatrix}.$$

The Hessian of f:

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial}{\partial x_1}(4x_1 - 2x_2) & \frac{\partial}{\partial x_2}(4x_1 - 2x_2) \\ \frac{\partial}{\partial x_1}(-2x_1 + 6x_2) & \frac{\partial}{\partial x_2}(-2x_1 + 6x_2) \end{pmatrix}$$

$$= \begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix}.$$

→ has eigenvalues $3.2679, 0.732 > 0$ $\therefore$ f is convex.

To find local min / max of f, set $\nabla f(x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and solve for x_1, x_2:

$$\begin{array}{l|l} 4x_1 - 2x_2 + 2 = 0 & 4x_1 - 2x_2 + 2 = 0 \\ -2x_1 + 6x_2 = 0 & \underline{-4x_1 + 12x_2 = 0 \ +} \\ & 10x_2 + 2 = 0. \end{array}$$

$\therefore$ $x_2 = -0.2$.

$\therefore$ $x_1 = \frac{\cancel{-}6 \cdot (-0.2)}{\cancel{-}2} = -0.6$.

$f(-0.6, -0.2) = 0.4$. $\quad \nabla f(0.6, -0.2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Q: How about <u>constrained</u> nonlinear optimization problems?
Do we have a good . condition for an optimal solution?

A: Yes!

<u>Ex</u>:

$$\left.\begin{array}{lll} \text{Minimize} & x_1^2 + x_2^2 & \\ \text{s.t.} & x_1^2 + x_2 \geq 4 & (1) \\ & x_1 + 4x_2 \geq 13. & (2) \end{array}\right\} (NLP)$$

Let
$$f(x) = x_1^2 + x_2^2.$$
$$g_1(x) = x_1^2 + x_2 - 4$$
$$g_2(x) = x_1 + 4x_2 - 13,$$

Then we want to solve:

$$\left.\begin{array}{ll} \text{Min} & f(x) \\ \text{s.t.} & g_1(x) \geq 0 \\ & g_2(x) \geq 0. \end{array}\right\} (NLP)$$

(*) As Q?

Since making sure that constraints are satisfied by (x_1, x_2) is "hard", let us consider an <u>unconstrained optimization</u> problem that is <u>related</u> to our original problem:
<u>Fix some $y_1 > 0$, $y_2 > 0$</u>: Then:
<u>Let</u> $L_y(x) := f(x) - y_1 g_1(x) - y_2 g_2(x)$.

Then: $\left.\text{Minimize} \quad L_y(x)\right\}$ (*)

I.e. Minimize $\underbrace{(x_1^2 + x_2^2)}_{\substack{\text{original} \\ \text{objective}}} \underbrace{- y_1 (x_1^2 + x_2 - 4) - y_2 (x_1 + 4x_2 - 13)}_{\text{"penalty terms"}}$

$$a = \lambda_1 a + \lambda_2$$
$$b = \lambda_1 + \lambda_2$$

<u>Remarks</u>: In the problem (*), x_1 and x_2 do not have to satisfy

① $x_1^2 + x_2 \geq 4$ and

② $x_1 + 4x_2 \geq 13$

However, if ① is not satisfied:

$x_1^2 + x_2 < 4$,

$\therefore \; -y_1(x_1^2 + x_2 - 4) > 0$

and if ② is not satisfied:

$x_1 + 4x_2 < 13$

$\therefore \; -y_2(x_1 + 4x_2 - 13) > 0$.

$\therefore$ we have to pay a "penalty" that increases the objective value, when the constraints are not satisfied.

~~In this example, it turns out that $L_y(x)$ is a convex function whenever $0 < y_1 < 1$.~~

So, we can solve it easily, by solving for x_1, x_2 such that

$$\nabla L_y(x) = 0.$$

<u>Ex</u>: $\nabla L(x) = \nabla f(x) - y_1 \nabla g_1(x) - y_2 \nabla g_2(x)$

$$= \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} - y_1 \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix} - y_2 \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

Q

$\therefore$ Solve: $\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} - y_1 \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix} - y_2 \begin{pmatrix} 1 \\ 4 \end{pmatrix}$.

for x_1 and x_2.

(it will be in terms of y_1 and y_2)

$$0 = 2x_1(1 - y_1) - y_2$$

$$0 = 2x_2 - (y_1 + 4y_2)$$

$$\rightarrow \quad x_1 = \frac{y_2}{2(1-y_1)} \qquad x_2 = \frac{y_1 + 4y_2}{2}.$$

Note, however, that for arbitrary choices of $y_1, y_2 \geq 0$ the resulting x_1, x_2 are not necessarily feasible or optimal for (NLP).

Ex: $y_1 = 0.5$, $y_2 = 0.5$

$$x_1 = \frac{0.5}{2(1+0.5)} = \frac{0.5}{1} = 0.5$$

$$x_2 = \frac{2.5}{2} = 1.25.$$

$x_1^2 + x_2 = 0.25 + 1.25 = 1.5 \not\geq 4$
$x_1 + 4x_2 = 0.5 + 5 = 5.5 \not\geq 13$.

$\therefore$ not feasible.

Ex: $y_1 = 0.5$ $y_2 = 10$.

$$x_1 = \frac{10}{2(0.5)} = 10$$

$$x_2 = \frac{0.5+40}{2} = 20.25$$

$x_1^2 + x_2 = 100 + 20.25 = 120.25 \geq 4$ ✓
$x_1 + 4x_2 = 10 + 81 = 91 \geq 13$ ✓. $\Big\}$ $\therefore$ feasible.

But $x_1^2 + x_2^2 = 100 + 410.0625 = 510.0625$ ← large
Q: is this optimal? maybe not.

Remark. • So far, it is not clear how solving for $L_y(x)$'s optimal solution gives us an opt solution for NLP, because we don't know what value of y to choose.
• However, if we are given a solution (x_1, x_2) feasible for (NLP), we can check whether it is opt or not by checking if there is an appropriate value of (y_1, y_2); ~~for which~~

(Necessary first order conditions of optimality)

Thm: Consider the following nonlinear program:

$$\left.\begin{array}{lll} \text{Min} & f(x) & \\ \text{s.t.} & g_i(x) \geq 0 & i=1,\dots,p \\ & h_j(x) = 0 & j=1,\dots,q. \end{array}\right\} (NLP)$$

(where f, g_i, h_j are continuous and diff'ble)

If $\bar{x}$ is an optimal solution for (NLP), then there exists Lagrange multipliers

$$\begin{array}{ll} y_i \geq 0 & i=1,\dots,p \\ z_j & j=1,\dots,q. \end{array}$$

That satisfy:

① if $g_i(\bar{x}) < 0$ then $y_i = 0$.

② $$\nabla_x L(x) = \nabla f(\bar{x}) - \sum_{i=1}^{p} y_i \nabla g_i(\bar{x}) - \sum_{j=1}^{q} z_j \nabla h_j(\bar{x}) = 0.$$

Remark: This thm presented conditions that are necessarily satisfied by an optimal solution $\bar{x}$ of (NLP).

EX: Consider $\bar{x} = (1,3)$ (I) Check that it is feasible for (NLP).

(II) To check if this is optimal, try to solve for y_1, y_2:

$$0 = 2(1) - y_1(2)(1) - y_2.$$
$$0 = 2(3) - y_1 - 4y_2.$$

$$0 = 2 - 2y_1 - y_2.$$
$$0 = 6 - y_1 - 4y_2.$$

$$0 = 2 - 2y_1 - y_2$$
$$0 = 12 - 2y_1 - 8y_2 \quad +$$

$$0 = -10 + 7y_2.$$

$$y_2 = \frac{10}{7} \qquad y_1 = 6 - 4\cdot\frac{10}{7} = \frac{2}{7}.$$

III Then, (y_1, y_2) satisfies 2.
Check if it satisfy 1 :

Since both $g_1(1,3) = 0$
$g_2(1,3) = 0$

then y_i can be $= 0$ or > 0. $\forall i = 1, 2$.

$\therefore$ 1 is satisfied.