ORIE 3310 / 5310

Lecture 25

23 April 2013

Recall from last time:

1. Necessary optimality condition for unconstrained optimization problem

$$\min_{x \in \mathbb{R}^n} f(x)$$

→ If $\bar{x}$ is an opt solution, then $\bar{x}$ satisfies: $\nabla f(\bar{x}) = \vec{0}$.

2. → If f is convex then:

$\bar{x}$ is an opt solution if and only if $\nabla f(\bar{x}) = \vec{0}$.

3. Necessary optimality condition for constrained opt problem

$$\begin{aligned} \min\ & f(x) \\ \text{s.t.}\ & g_i(x) \geq 0 \quad i = 1, \dots, p \\ & h_j(x) = 0 \quad j = 1, \dots, q. \end{aligned}$$

→ If $\bar{x}$ is an opt solution, then $\bar{x}$ satisfies:

There exists $y_i \geq 0 \quad i = 1, \dots, p$

$z_j \quad j = 1, \dots, q$,

Such that

1. if $g_i(\bar{x}) > 0$, then $y_i = 0$
2. $\nabla L_{y,z}(\bar{x}) = 0$

where $L_{y,z}(x) = f(x) - \sum_1^p y_i g_i(x) - \sum_1^q z_j h_j(x)$

(so, $\nabla L_{y,z}(x) = \nabla f(x) - \sum_{i=1}^p y_i \nabla g_i(x) - \sum_{j=1}^q z_j \nabla h_j(x)$)

④ We can use ③ to check if a proposed $\bar{x}$ is opt for (NLP) or not :

<u>Step 1</u> : Check that $\bar{x}$ is feasible.
That is, that $g_i(\bar{x}) \geq 0 \ \forall i = 1, \dots, p$.
$h_j(\bar{x}) = 0 \ \forall j = 1, \dots, q$.

<u>Step 3</u> : If $g_i(\bar{x}) > 0$, let $y_i = 0$.

<u>Step 2</u> : Compute $\nabla L_{y,z}(\bar{x})$.
~~For the~~ Plug in y_i's that are zeros

Solve $\nabla L_{y,z}(\bar{x}) = 0$ for y, z.

If there ~~exi~~ are y_i's, z_j's that satisfy steps 2, 3, then $\bar{x}$ ~~is~~ ~~cannot~~ might be opt. otherwise, $\bar{x}$ is not optimal.

=.

<u>Today</u> : ① Some geometric Interpretation
② Newton's method.

=.

<u>Geometric Interpretation</u> : [illegible].
of the Necessary conditions for Constrained Problems :

Ⓐ First, consider Problems with just equality constraints :

$$\begin{aligned} &\text{Min } f(x) \\ &\text{s.t. } h(x) = 0. \end{aligned}$$

<u>Ex</u> :

$$\begin{aligned} &\text{Min } x_1 + x_2 \\ &\text{s.t. } x_1^2 + x_2^2 = 4. = 0. \end{aligned} \quad \Bigg| \quad \begin{aligned} &f(x) = x_1 + x_2. \\ &h_1(x) = x_1^2 + x_2^2 - 4. \end{aligned}$$

If $\bar{x}$ is optimal, then $\exists\, z$ such that

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

$$L_z(x) = f(x) - z h(x)$$

$$\nabla L_z(x) = \nabla f(x) - z \nabla h(x).$$

$$= \begin{pmatrix}1\\1\end{pmatrix} - z\begin{pmatrix}2x_1\\2x_2\end{pmatrix}.$$

$$\nabla L_z(x) = 0$$

$$\therefore \begin{pmatrix}1\\1\end{pmatrix} - z\begin{pmatrix}2x_1\\2x_2\end{pmatrix} = 0.$$

$$\begin{pmatrix}1\\1\end{pmatrix} = z\begin{pmatrix}2x_1\\2x_2\end{pmatrix}. \Rightarrow \left.\begin{array}{l}1 = z\cdot 2x_1\\ 1 = z\cdot 2x_2\end{array}\right\} z = \frac{1}{2x_1} = \frac{1}{2x_2}.$$

$\therefore$ there exists such z if and only if

- $\bar{x}$ is feasible, and.
- $\bar{x}_1 = \bar{x}_2$.

① $\bar{x}$ feasible means: $x_1^2 + x_2^2 = 4$.

② $\bar{x}_1 = \bar{x}_2$ $\therefore$ $x_1^2 + x_2^2 = 2x_1^2 = 4$.

$$x_1^2 = 2.$$

$$x_1 = \pm\sqrt{2}$$

$\therefore$ $(x_1, x_2) = (\sqrt{2}, \sqrt{2})$ or $(x_1, x_2) = (-\sqrt{2}, -\sqrt{2})$.

$\rightarrow z = \frac{1}{\sqrt{2}}$ $\qquad \rightarrow z = -\frac{1}{\sqrt{2}}$.

$$\nabla^2 L_z(x) = -z\begin{pmatrix}2 & 0\\0 & 2\end{pmatrix}.$$

at $z = \frac{1}{\sqrt{2}}$, $\nabla^2 L_z(x) = -\frac{1}{\sqrt{2}}\begin{pmatrix}2 & 0\\0 & 2\end{pmatrix} = \begin{pmatrix}-\sqrt{2} & 0\\0 & -\sqrt{2}\end{pmatrix}$

at $z = -\frac{1}{\sqrt{2}}$, $\nabla^2 L_z(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}2 & 0\\0 & 2\end{pmatrix} = \begin{pmatrix}\sqrt{2} & 0\\0 & \sqrt{2}\end{pmatrix}$.

$\therefore$ $x = (-\sqrt{2}, -\sqrt{2})$ is the <u>minimizer</u>.

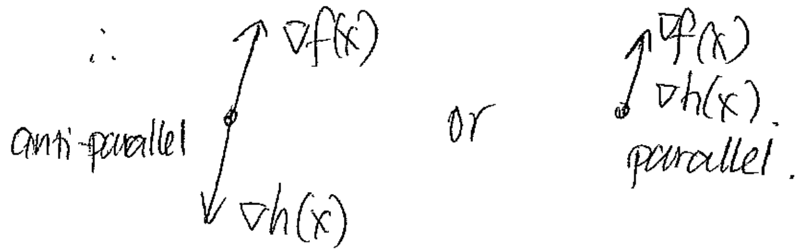
In picture, the condition that

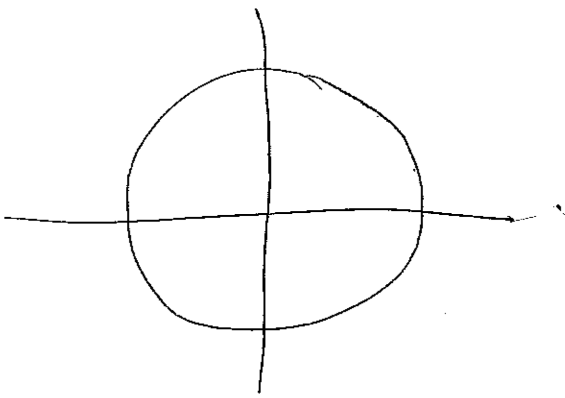
$$\nabla L_{z}(x) = 0$$

is: $\nabla f(x) = z \nabla h(x)$.

$\therefore$ $\nabla f(x)$ is a scalar multiple of $\nabla h(x)$.

$\therefore$





$\nabla f(x)$ = direction of greatest increase of f.

$\therefore$ $-\nabla f(x)$ = direction of greatest decrease of f.

$\nabla h(x)$ = normal to the space $h(x)$

i.e. if we go in direction $\nabla h(x)$, will leave the feasible region.

$\therefore$ the condition $\nabla f(x) = z \nabla h(x)$ says that the direction in which you can decrease $f(x)$ is not a feasible direction.

(B) Consider problems with inequality constraints.

The conditions
(1) If $g_i(x) > 0$ then $y_i = 0$
(2) $\nabla L_y(x) = 0$

Says that:

- If the optimal solution $\bar{x}$ is on the boundary of the ~~feas region~~ i^{th} constraint
  i.e. $g_i(x) = 0$.
  then treat $g_i(x)$ like an equality constraint.
  $\rightarrow$ like in case (A).
- If the optimal solution $\bar{x}$ is not on the boundary of the i^{th} constraint
  i.e. $g_i(x) > 0$,
  then we can "pretend" that the constraint $g_i(x)$ is not present.

Newton's Method.

(A) General Discussion.

- Newton's Method is an algorithm/method for numerically finding roots of an equation.

For example:

$$\phi(x) = x^2 - 3x + 2.$$

Solve for x such that $\phi(x) = 0$.

$$x^2 - 3x + 2 = 0.$$

$$x = \frac{3 \pm \sqrt{9-8}}{2} \quad \text{quadratic formula.}$$

$$x = \frac{3+1}{2}, \quad x = \frac{3-1}{2}$$

$$\therefore x = 2, \quad x = 1.$$

Solving this problem is "easy" because $\varphi(x)$ happens to be a quadratic function.

However, in general, it is hard to solve it directly, E.g. if $\varphi(x)$ is a polynomial of high degree, or if it is a nonalgebraic function.

<u>Ex</u>: $\varphi(x) = x^6 - 5x^5 + 0.3x^4 + x^3 + 4x^2 - 5x + 7$.

$\rightarrow$ No closed form solution, in general.

Newton's Method provide a reliable way to find a root of such equations.

~~Idea: The gradient of f~~

<u>Method</u>:

① Start at a point $x^{(0)}$.

② At the i^{th} iteration: $x^{(i)}$ is the current point

- Compute $\varphi'(x^{(i)})$.
- Compute $\varphi(x^{(i)})$.
- Let $\delta = -\varphi(x^{(i)}) / \varphi'(x^{(i)})$.
- Let $x^{(i+1)} = x^{(i)} + \delta$.

<u>Explanation</u>:

- Suppose we are at the point $x^{(0)}$
- The best linear approximation of φ at $x^{(0)}$ is:

$$\tilde{\varphi}(x^{(0)} + \delta) = \varphi(x^{(0)}) + \delta \varphi'(x^{(0)}).$$

- Solve for δ such that

$$\tilde{\varphi}(x^{(0)} + \delta) = 0.$$

$\therefore \ \phi(x^{(0)}) + \delta\phi'(x^{(0)}) = 0 .$

$$\delta\phi'(x^{(0)}) = -\phi(x^{(0)})$$

$$\delta = -\frac{\phi(x^{(0)})}{\phi'(x^{(0)})} .$$

$\therefore$ ~~Approx~~ The point $\tilde{x}$ that achieves

$$\hat{\phi}(\tilde{x}) = 0$$

is $\tilde{x} = x^{(0)} + \delta$.

$$\therefore \text{ let } x^{(1)} = x^{(0)} + \delta .$$

<u>Remarks</u>: