

6.252 NONLINEAR PROGRAMMING

LECTURE 11

CONSTRAINED OPTIMIZATION;

LAGRANGE MULTIPLIERS

LECTURE OUTLINE

- Equality Constrained Problems
- Basic Lagrange Multiplier Theorem
- Proof 1: Elimination Approach
- Proof 2: Penalty Approach

Equality constrained problem

minimize $f(x)$

subject to $h_i(x) = 0, \quad i = 1, \dots, m.$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$, $h_i : \mathbb{R}^n \mapsto \mathbb{R}$, $i = 1, \dots, m$, are continuously differentiable functions. (Theory also applies to case where f and h_i are cont. differentiable in a neighborhood of a local minimum.)

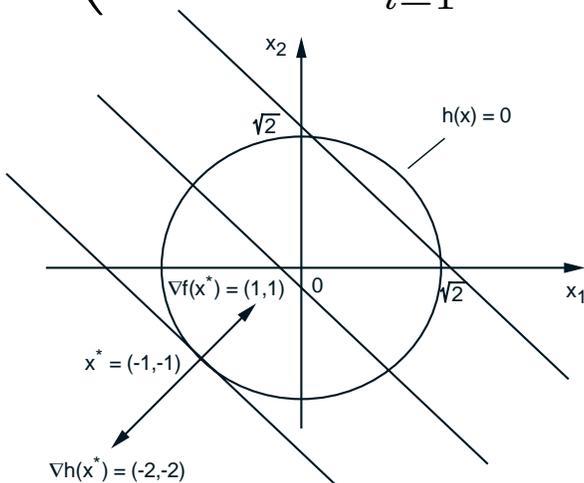
LAGRANGE MULTIPLIER THEOREM

- Let x^* be a local min and a *regular point* [$\nabla h_i(x^*)$: linearly independent]. Then there exist unique scalars $\lambda_1^*, \dots, \lambda_m^*$ such that

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla h_i(x^*) = 0.$$

If in addition f and h are twice cont. differentiable,

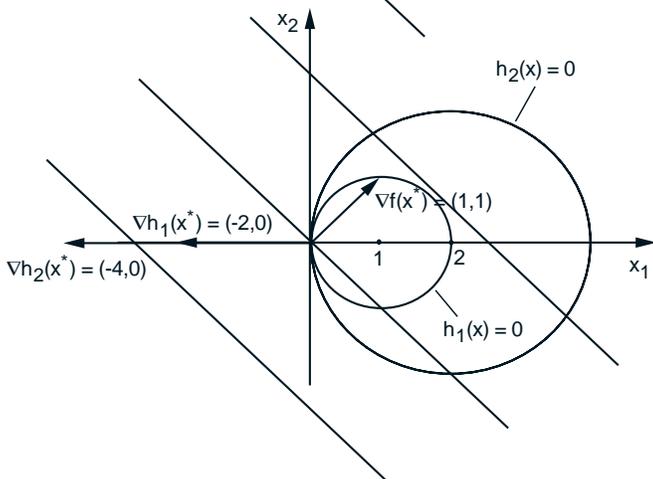
$$y' \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) \right) y \geq 0, \quad \forall y \text{ s.t. } \nabla h(x^*)' y = 0$$



minimize $x_1 + x_2$

subject to $x_1^2 + x_2^2 = 2$.

The Lagrange multiplier is $\lambda = 1/2$.



minimize $x_1 + x_2$

s. t. $(x_1 - 1)^2 + x_2^2 - 1 = 0$

$(x_1 - 2)^2 + x_2^2 - 4 = 0$

PROOF VIA ELIMINATION APPROACH

- Consider the linear constraints case

$$\text{minimize } f(x)$$

$$\text{subject to } Ax = b$$

where A is an $m \times n$ matrix with linearly independent rows and $b \in \mathfrak{R}^m$ is a given vector.

- Partition $A = \begin{pmatrix} B & R \end{pmatrix}$, where B is $m \times m$ invertible, and $x = \begin{pmatrix} x_B & x_R \end{pmatrix}'$. Equivalent problem:

$$\text{minimize } F(x_R) \equiv f\left(B^{-1}(b - Rx_R), x_R\right)$$

$$\text{subject to } x_R \in \mathfrak{R}^{n-m}.$$

- Unconstrained optimality condition:

$$0 = \nabla F(x_R^*) = -R'(B')^{-1} \nabla_B f(x^*) + \nabla_R f(x^*) \quad (1)$$

By defining

$$\lambda^* = -(B')^{-1} \nabla_B f(x^*),$$

we have $\nabla_B f(x^*) + B' \lambda^* = 0$, while Eq. (1) is written $\nabla_R f(x^*) + R' \lambda^* = 0$. Combining:

$$\nabla f(x^*) + A' \lambda^* = 0$$

ELIMINATION APPROACH - CONTINUED

- Second order condition: For all $d \in \mathfrak{R}^{n-m}$

$$0 \leq d' \nabla^2 F(x_R^*) d = d' \nabla^2 \left(f \left(B^{-1}(b - Rx_R), x_R \right) \right) d. \quad (2)$$

- After calculation we obtain

$$\begin{aligned} \nabla^2 F(x_R^*) &= R'(B')^{-1} \nabla_{BB}^2 f(x^*) B^{-1} R \\ &\quad - R'(B')^{-1} \nabla_{BR}^2 f(x^*) - \nabla_{RB}^2 f(x^*) B^{-1} R + \nabla_{RR}^2 f(x^*). \end{aligned}$$

- Eq. (2) and the linearity of the constraints [implying that $\nabla^2 h_i(x^*) = 0$], yields for all $d \in \mathfrak{R}^{n-m}$

$$\begin{aligned} 0 &\leq d' \nabla^2 F(x_R^*) d = y' \nabla^2 f(x^*) y \\ &= y' \left(\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*) \right) y, \end{aligned}$$

where $y = (y_B \quad y_R)' = (-B^{-1}Rd \quad d)'$.

- y has this form iff

$$0 = By_B + Ry_R = \nabla h(x^*)' y.$$

PROOF VIA PENALTY APPROACH

- Introduce, for $k = 1, 2, \dots$, the cost function

$$F^k(x) = f(x) + \frac{k}{2} \|h(x)\|^2 + \frac{\alpha}{2} \|x - x^*\|^2,$$

where $\alpha > 0$ and x^* is a local minimum.

- Let $\epsilon > 0$ be such that $f(x^*) \leq f(x)$ for all feasible x in the *closed* sphere $S = \{x \mid \|x - x^*\| \leq \epsilon\}$, and let $x^k = \arg \min_{x \in S} F^k(x)$. Have

$$F^k(x^k) = f(x^k) + \frac{k}{2} \|h(x^k)\|^2 + \frac{\alpha}{2} \|x^k - x^*\|^2 \leq F^k(x^*) = f(x^*)$$

Hence, $\lim_{k \rightarrow \infty} \|h(x^k)\| = 0$, so for every limit point \bar{x} of $\{x^k\}$, $h(\bar{x}) = 0$.

- Furthermore, $f(x^k) + (\alpha/2) \|x^k - x^*\|^2 \leq f(x^*)$ for all k , so by taking lim,

$$f(\bar{x}) + \frac{\alpha}{2} \|\bar{x} - x^*\|^2 \leq f(x^*).$$

Combine with $f(x^*) \leq f(\bar{x})$ [since $\bar{x} \in S$ and $h(\bar{x}) = 0$] to obtain $\|\bar{x} - x^*\| = 0$ so that $\bar{x} = x^*$. Thus $\{x^k\} \rightarrow x^*$.

PENALTY APPROACH - CONTINUED

- Since $x^k \rightarrow x^*$, for large k , x^k is interior to S , and is an *unconstrained* local minimum of $F^k(x)$.
- From 1st order necessary condition,

$$0 = \nabla F^k(x^k) = \nabla f(x^k) + k \nabla h(x^k) h(x^k) + \alpha(x^k - x^*). \quad (3)$$

Since $\nabla h(x^*)$ has rank m , $\nabla h(x^k)$ also has rank m for large k , so $\nabla h(x^k)' \nabla h(x^k)$: invertible. Thus, multiplying Eq. (3) w/ $\nabla h(x^k)'$

$$kh(x^k) = - \left(\nabla h(x^k)' \nabla h(x^k) \right)^{-1} \nabla h(x^k)' \left(\nabla f(x^k) + \alpha(x^k - x^*) \right).$$

Taking limit as $k \rightarrow \infty$ and $x^k \rightarrow x^*$,

$$\{ kh(x^k) \} \rightarrow - \left(\nabla h(x^*)' \nabla h(x^*) \right)^{-1} \nabla h(x^*)' \nabla f(x^*) \equiv \lambda^*.$$

Taking limit as $k \rightarrow \infty$ in Eq. (3), we obtain

$$\nabla f(x^*) + \nabla h(x^*) \lambda^* = 0.$$

- 2nd order L-multiplier condition: Use 2nd order unconstrained condition for x^k , and algebra.

LAGRANGIAN FUNCTION

- Define the Lagrangian function

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x).$$

Then, if x^* is a local minimum which is regular, the Lagrange multiplier conditions are written

$$\nabla_x L(x^*, \lambda^*) = 0, \quad \nabla_\lambda L(x^*, \lambda^*) = 0,$$

System of $n + m$ equations with $n + m$ unknowns.

$$y' \nabla_{xx}^2 L(x^*, \lambda^*) y \geq 0, \quad \forall y \text{ s.t. } \nabla h(x^*)' y = 0.$$

- Example

$$\text{minimize } \frac{1}{2} (x_1^2 + x_2^2 + x_3^2)$$

$$\text{subject to } x_1 + x_2 + x_3 = 3.$$

Necessary conditions

$$x_1^* + \lambda^* = 0, \quad x_2^* + \lambda^* = 0,$$

$$x_3^* + \lambda^* = 0, \quad x_1^* + x_2^* + x_3^* = 3.$$

EXAMPLE - PORTFOLIO SELECTION

- Investment of 1 unit of wealth among n assets with random rates of return e_i , and given means \bar{e}_i , and covariance matrix $Q = [E\{(e_i - \bar{e}_i)(e_j - \bar{e}_j)\}]$.

- If x_i : amount invested in asset i , we want to

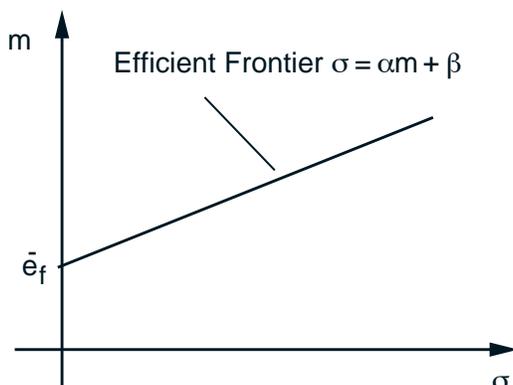
minimize Variance of return $y = \sum_i e_i x_i = x' Q x$

subject to $\sum_i x_i = 1$, and given mean $\sum_i \bar{e}_i x_i = m$

- Let λ_1 and λ_2 be the L-multipliers. Have $2Qx^* + \lambda_1 u + \lambda_2 \bar{e} = 0$, where $u = (1, \dots, 1)'$ and $\bar{e} = (\bar{e}_1, \dots, \bar{e}_n)'$. This yields

$$x^* = mv + w, \quad \sigma^2 = (\alpha m + \beta)^2 + \gamma,$$

where v and w are vectors, and α , β , and γ are some scalars that depend on Q and \bar{e} .



If riskless asset exists, $\gamma = 0$. Efficient frontier is a line (CAPM model).