Optimization. A first course of mathematics for economists

Xavier Martinez-Giralt
Universitat Autònoma de Barcelona
xavier.martinez.giralt@uab.eu

II.2 Static optimization - Classical programming

Classical programming

Problem:

- $\max_{\mathbf{x}} f(\mathbf{x})$ s.t. $\mathbf{g}(\mathbf{x}) = \mathbf{b}$

where:

- $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbf{R}^n$
- $\mathbf{g}(\mathbf{x}) = \begin{pmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{pmatrix}; \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$
- $g_i(\mathbf{x}) = b_i, \; i = 1, \ldots, m$
- $g_i(\mathbf{x})$ continuous, continuously differentiable
- $b_i \in \mathbf{R}, \; \mathbf{x} \in \mathbf{R}^n$

Classical programming (2)

Feasible set:

- $X = \{\mathbf{x} \in \mathbf{R}^n | \mathbf{g}(\mathbf{x}) = \mathbf{b}\}$. Points $\mathbf{x} \in \mathbf{R}^n \cap_{i=1}^{m} g_i(\mathbf{x})$
- Problem: find the set of points in X in the highest level set of objective function

Three possibilities:

- $n > m$. The difference $n - m$ is the *degrees of freedom* of the problem.
- $n = m$. Problem is trivial. Consider: $\max_x ax^2$, s.t. $bx = c$. Equivalently $\max_x a(\frac{c}{d})^2$ that does not depend of x.
- $n < m$. Either there are $m - n$ redundant restrictions, or restrictions are inconsistent among them, and set of solutions is empty.

A particular case $m = 0$. Free maximization

Getting the intuition with $n = 1$:

- Let $f : X \to \mathbf{R}$ be twice continuously differentiable on $X \subset \mathbf{R}$.
- Find $x^* \in \mathbf{R}$ solution of $\max_x f(x)$; Let Δx be arbitrarily small.
- First order (necessary) condition
  - Given that $x^* \in \mathbf{R}$ is solution, $f(x^*) \geq f(x^* + \Delta x)$
  - Taylor expansion around x^*: (with $\theta \in (0,1)$)
    $f(x^* + \Delta x) = f(x^*) + \frac{df}{dx}(x^*)\Delta x + \frac{1}{2}\frac{d^2 f}{dx^2}(x^* + \theta\Delta x)(\Delta x)^2$
  - Define *fundamental inequality* $(FI) \equiv f(x^* + \Delta x) - f(x^*)$:
    $FI(\Delta x) \equiv \Delta x[\frac{df}{dx}(x^*) + \frac{1}{2}\frac{d^2 f}{dx^2}(x^* + \theta\Delta x)(\Delta x)] \leq 0$
  - If $\Delta x > 0$, then $\lim_{\Delta x \to 0} FI(\Delta x) = \frac{df}{dx}(x^*) \leq 0$
  - If $\Delta x < 0$, then $\lim_{\Delta x \to 0} FI(\Delta x) = \frac{df}{dx}(x^*) \geq 0$
- Hence, FI implies as FOnC to maximize $f(x)$ that $\frac{df}{dx}(x^*) = 0$.

A particular case $m = 0$. Free maximization (2)

Getting the intuition with $n = 1$:

- Second order (necessary) condition
  - Substituting FOC in FI we obtain $\frac{d^2 f}{dx^2}(x^* + \theta \Delta x) \leq 0$.
  - As it is verified $\forall \Delta x$, it follows that SOnC is $\frac{d^2 f}{dx^2}(x^*) \leq 0$.
- Sufficient conditions for a local maximum: If

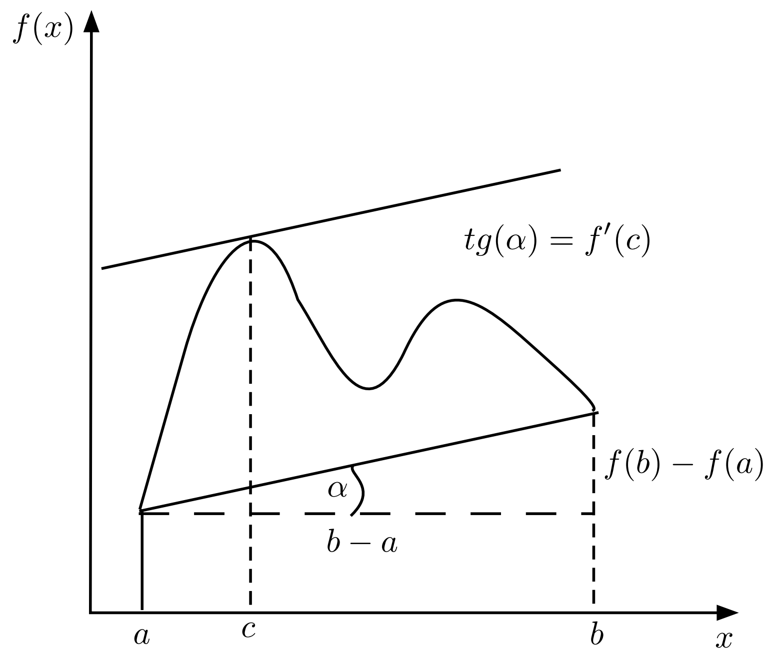
$$\frac{df}{dx}(x^*) = 0, \text{ and}$$
$$\frac{d^2 f}{dx^2}(x^*) < 0,$$

then, $f(x^*) > f(x^* + \Delta x)$.
Proof: use either *mean value theorem* or FI.

A particular case $m = 0$. Free maximization (3)

Getting the intuition with $n = 1$:

- Mean value theorem
  - Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable on $(a, b),\ a < b$.
  - Then, $\exists c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

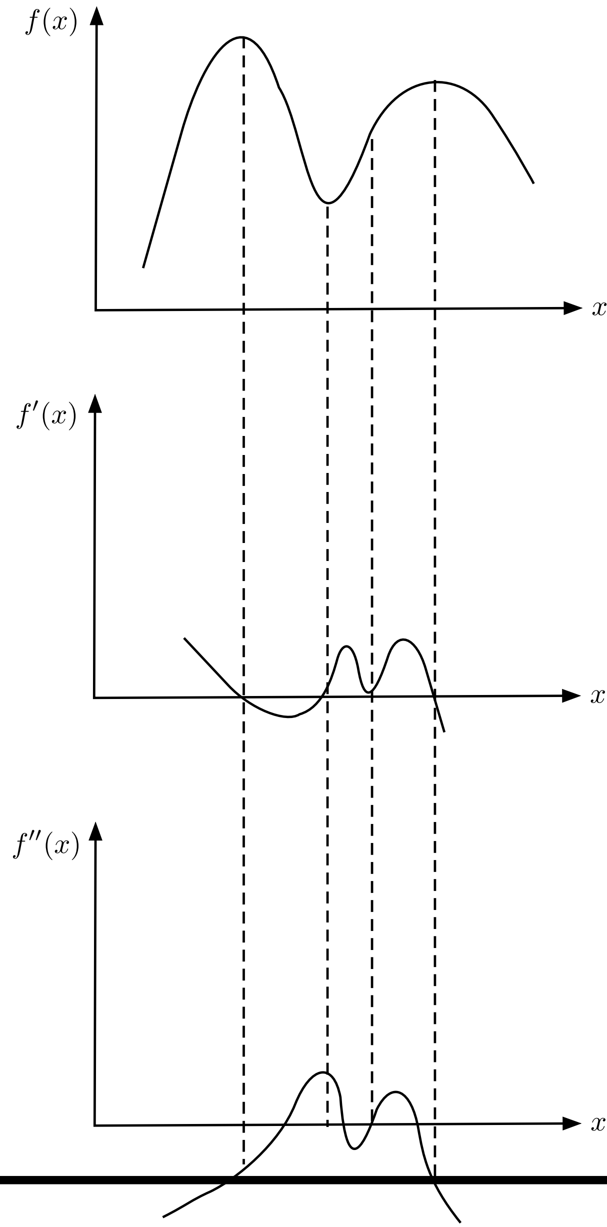


A particular case $m = 0$. Free maximization (3)

Getting the intuition with $n = 1$:

- Sufficient conditions - Proof
  - Let $a = x^*$, let $b = x^* + \Delta x$, and let $c = x^* + \theta \Delta x,\ \theta \in (0,1)$.
  - Mean value theorem says: $\frac{df}{dx}(x^* + \theta\Delta x) = \frac{f(x^*+\Delta x) - f(x^*)}{(x^*+\Delta x) - x^*} = \frac{f(x^*+\Delta x) - f(x^*)}{\Delta x}$
  - or equivalently $f(x^* + \Delta x) = f(x^*) + \frac{df}{dx}(x^* + \theta\Delta x)\Delta x$
  - given that $f'(x^*) = 0$ and $f''(x^*) < 0$, for $\Delta x > 0$, necessarily $\frac{df}{dx}(x^* + \theta\Delta x) < 0$.
  - given that $f'(x^*) = 0$ and $f''(x^*) < 0$, for $\Delta x < 0$, necessarily $\frac{df}{dx}(x^* + \theta\Delta x) > 0$.
  - hence, $\frac{df}{dx}(x^* + \theta\Delta x)\Delta x < 0$ and thus $f(x^* + \Delta x) < f(x^*)$.

Geometry of free maximization in $\mathbb{R}$



Free maximization with $n > 1$

- Problem: $\max_{x_1,\ldots,x_n} f(x_1,\ldots,x_n)$
- Theorem: (proof parallel to the case $n = 1$)
  - Let $f : X \to \mathbf{R}$ be twice continuously differentiable on $X \subset \mathbf{R}^n$.
  - Let $\mathbf{x}^* \in \mathbf{R}^n$ be a local maximum of f.
  - Then, FOnCs are $\frac{\partial f(x)}{\partial x_j}(\mathbf{x}^*) = 0, \ \forall j.$
- SOnC : Hessian negative semidefinite: $(\Delta\mathbf{x}^*)'\frac{\partial^2 f}{\partial x^2}(\mathbf{x}^*)(\Delta\mathbf{x}^*) \leq 0, \ \forall \mathbf{x}$
- Sufficient conditions:

$$\frac{\partial f(x)}{\partial x_j}(\mathbf{x}^*) = 0, \ \forall j.$$

$$(\Delta\mathbf{x}^*)'\frac{\partial^2 f}{\partial x^2}(\mathbf{x}^*)(\Delta\mathbf{x}^*) < 0, \ \forall \mathbf{x}$$

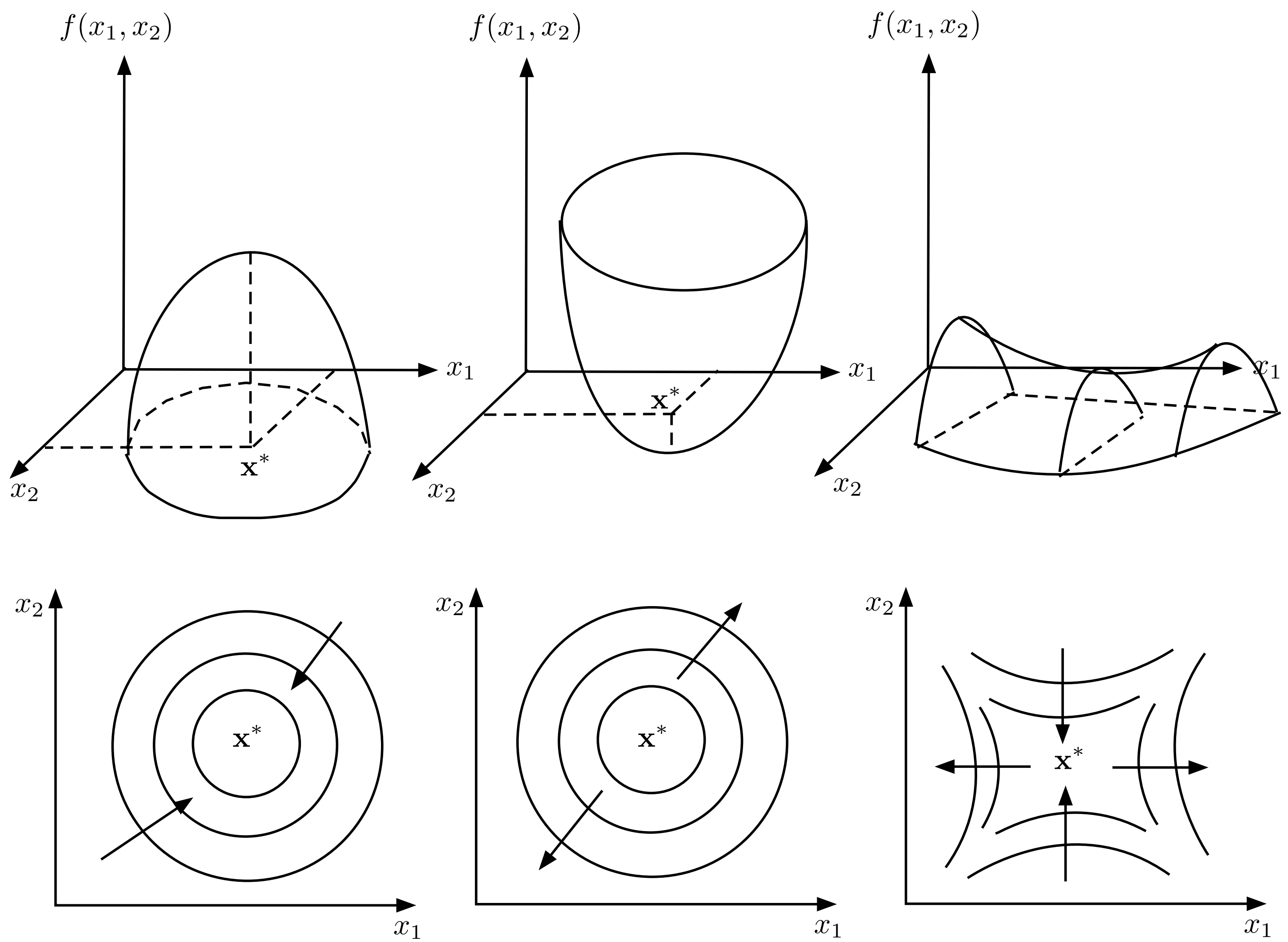
Free maximization. Illustration with $n = 2$

- Problem: $\max_{x_1, x_2} f(x_1, x_2)$
- Let $(\mathbf{x}^*)' = (x_1^*, x_2^*)'$ be a local maximum.
- FOnC: $\frac{\partial f}{\partial x_1}(\mathbf{x}^*) = 0, \ \frac{\partial f}{\partial x_2}(\mathbf{x}^*) = 0$
- SOnC: Hessian negative semidefinite, i.e.

$$\frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^*) \leq 0,$$

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^*) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^*) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^*) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}^*) \end{vmatrix} \geq 0$$

Geometry of Free maximization with $n = 2$



Classical programming (3)

Two strategies

- Substitute restrictions $\rightarrow$ solve a free maximization problem
- General method: Lagrange multipliers

Method of substitution. Illustration

- Problem: $\max_{x_1, x_2} f(x_1, x_2)$ s.t. $g(x_1, x_2) = 0$
- Assuming $\frac{\partial g}{\partial x_2} \neq 0$, we have an implicit function $x_2(x_1)$.
- Using the implicit function theorem we know, $\frac{\partial x_2}{\partial x_1} = -\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}}$
- Problem: $\max_{x_1} f(x_1, x_2(x_1))$ with FOC: $\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial x_1} = 0$
- and the solution is: $\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} = \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}}, \; g(x_1, x_2) = 0$
- i.e. set of tangency points between f function and g function.

Classical programming (4)

Method of substitution. Illustration (cont'd)

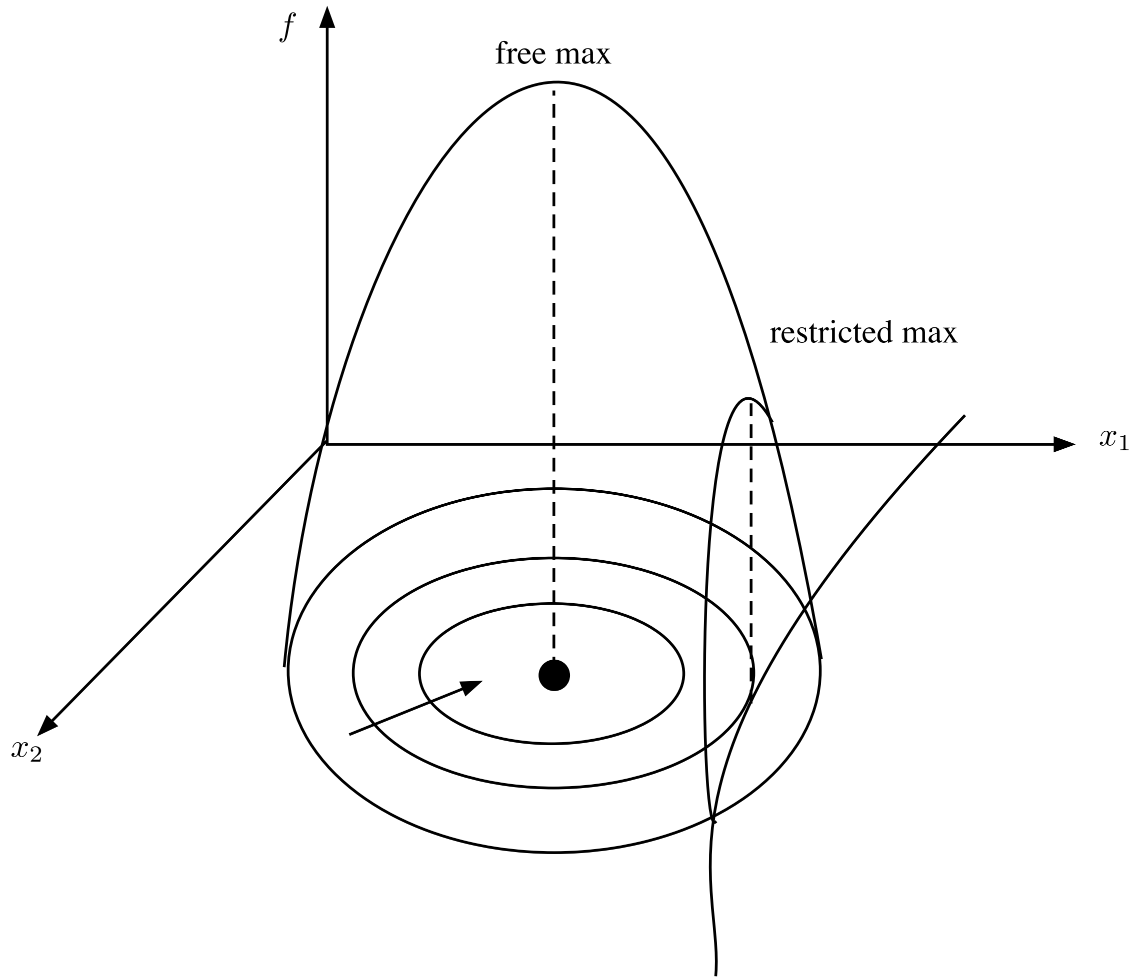
- Rewrite FOC as (for future use):

$$\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2} \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} = 0$$

$$\frac{\partial f}{\partial x_1} + \lambda \frac{\partial g}{\partial x_1} = 0$$

with $\lambda = -\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}}$

Geometry of the classical programming



Classical programming. Lagrange method.

Lagrange theorem:

- Let $f, g_i,\ i = 1, \dots m$ be twice continuously differentiable.
- Suppose $\mathbf{x}^* = (x_1^*, \dots x_n^*)$ is a local interior extreme point of $f(\mathbf{x}^*)$ subject to $g_i(x_1, \dots x_n) = b_i,\ i = 1, \dots, m$.
- Suppose $Jg(\mathbf{x}^*) \neq 0$.
- Then, $\exists \lambda^* = (\lambda_1^*, \dots, \lambda_n^*)$ such that $(\mathbf{x}^*, \lambda^*)$ is a critical point of the lagrangean function $L(\mathbf{x}, \lambda) = f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i (b_i - g_i(\mathbf{x}^*))$.
- where λ_i gives the (shadow) price associated with constraint i.
- Equivalently, $\exists \lambda^*$ such that $\nabla f(\mathbf{x}^*) = \lambda^* \nabla g(\mathbf{x}^*)$

Classical programming. Lagrange method (2)

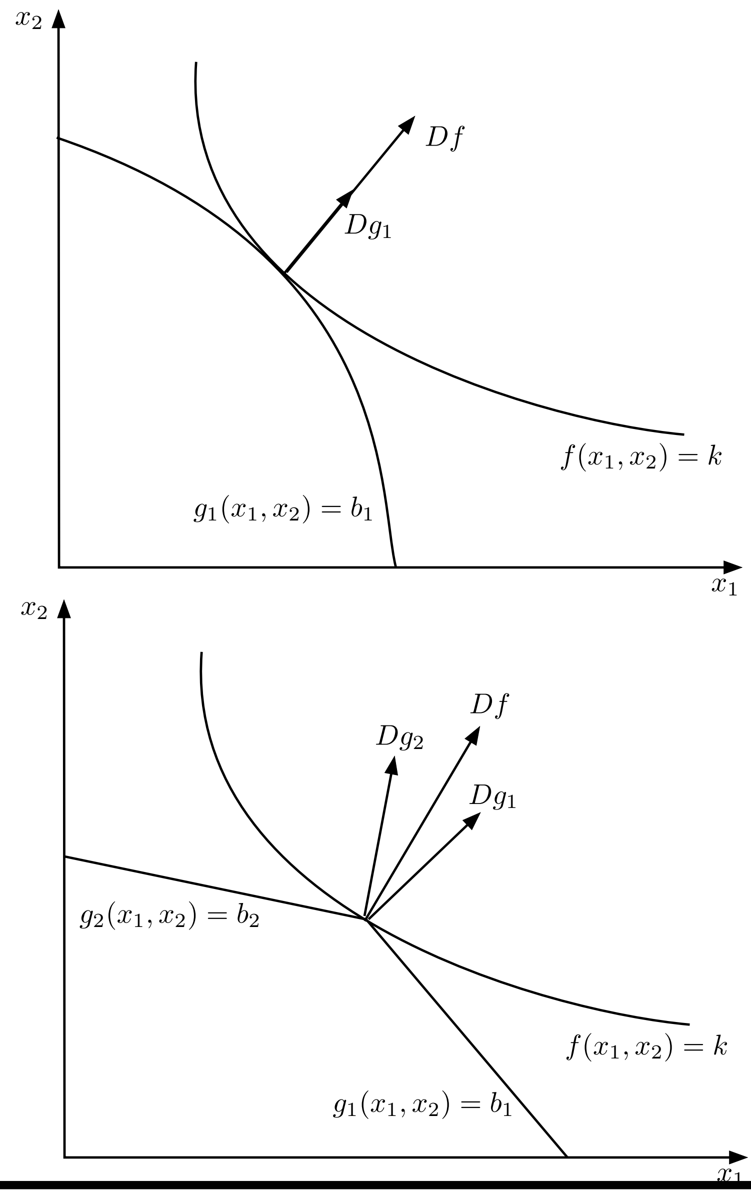
Lagrange theorem - Remarks

- $Jg(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial g_1(\mathbf{x}^*)}{\partial x_1} & \cdots & \frac{\partial g_1(\mathbf{x}^*)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m(\mathbf{x}^*)}{\partial x_1} & \cdots & \frac{\partial g_m(\mathbf{x}^*)}{\partial x_n} \end{pmatrix} \neq 0$ means that the matrix has rank m or that the m restrictions are linearly independent.
- This is known as the constraint qualification.
- $\lambda_i^* \lesseqgtr 0$
- λ_i^* satisfies $\frac{\partial f(\mathbf{x}^*)}{\partial x_j} = \sum_{i=1}^m \lambda_i^* \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j}$, $j = 1, \ldots, n$. i.e.
  - gradient vector of objective function $=$ sum of product of lagrange multiplier by Jacobian of restrictions, or
  - gradient vector of objective function is a linear combination of gradients of restrictions.

Lagrange method - Interpretation

- Note that the FOC of lagrangrean function is the same as the FOC of the problem under the substitution method.
- Graphically the gradient of the objective function must be contained in the cone formed by the gradients of the restrictions.
- Otherwise, we could achieve a higher value of the objective function without violating the restrictions. Contradiction with the assumption of $\mathbf{x}^*$ being a local maximum.
- Once we have found candidate solutions $\mathbf{x}^*$, it is not always easy to assess whether they correspond to a minimum, a maximum or neither. (FOCs are necessary conditions). Two particular cases:
  - If f concave and g_i linear, then $\mathbf{x}^*$ are local maxima.
  - If f convex and g_i linear, then $\mathbf{x}^*$ are local minima.

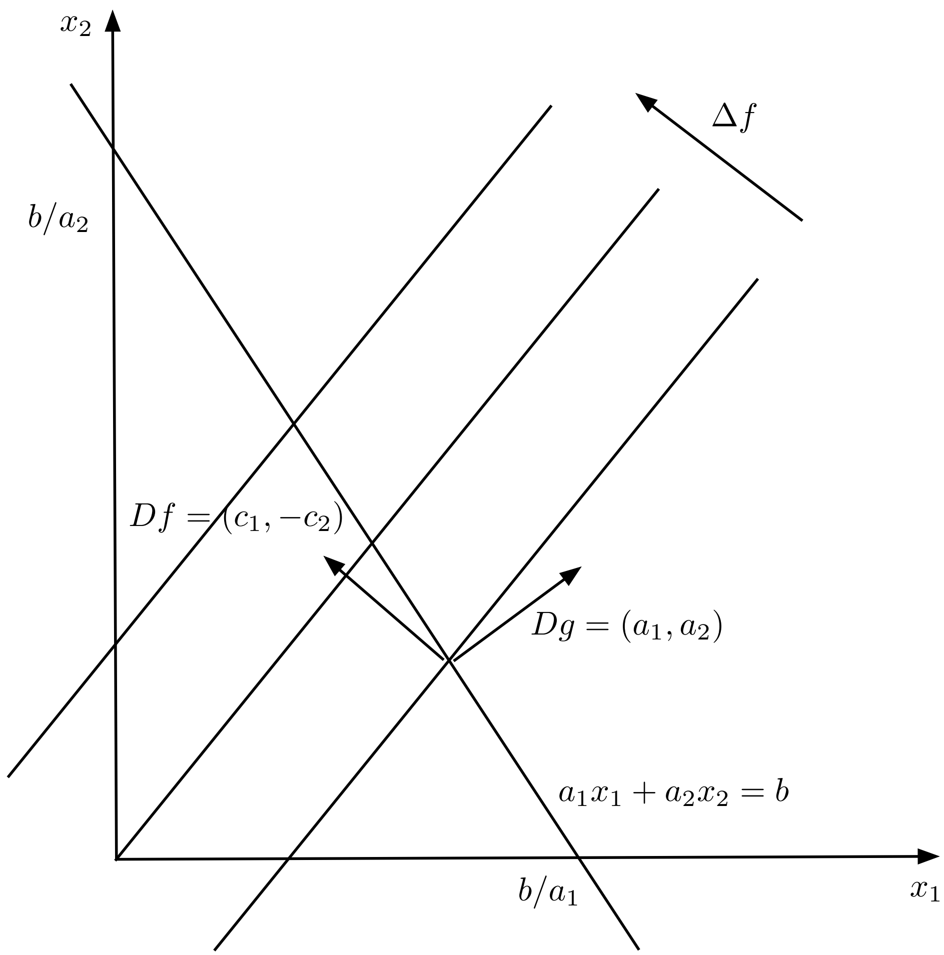
Geometry of the constraint qualification



An example violating the constraint qualification

- $n = 2,\ m = 1$
- $\max_{x_1,x_2} c_1x_1 - c_2x_2$ s.t. $a_1x_1 + a_2x_2 = b,\ c_i, a_i > 0$
- $L(x_1, x_2, \lambda) = c_1x_1 - c_2x_2 + \lambda(b - a_1x_1 - a_2x_2)$
- $\frac{\partial L}{\partial x_1} = c_1 - \lambda a_1 = 0$
- $\frac{\partial L}{\partial x_2} = -c_2 - \lambda a_2 = 0$
- $\frac{\partial L}{\partial \lambda} = b - a_1x_1 - a_2x_2 = 0$
- from the first two conditions we obtain $\lambda = \frac{c_1}{a_1};\ \lambda = \frac{-c_2}{a_2}$
- Contradiction!! $\rightarrow$ no solution.
- Graphical representation:

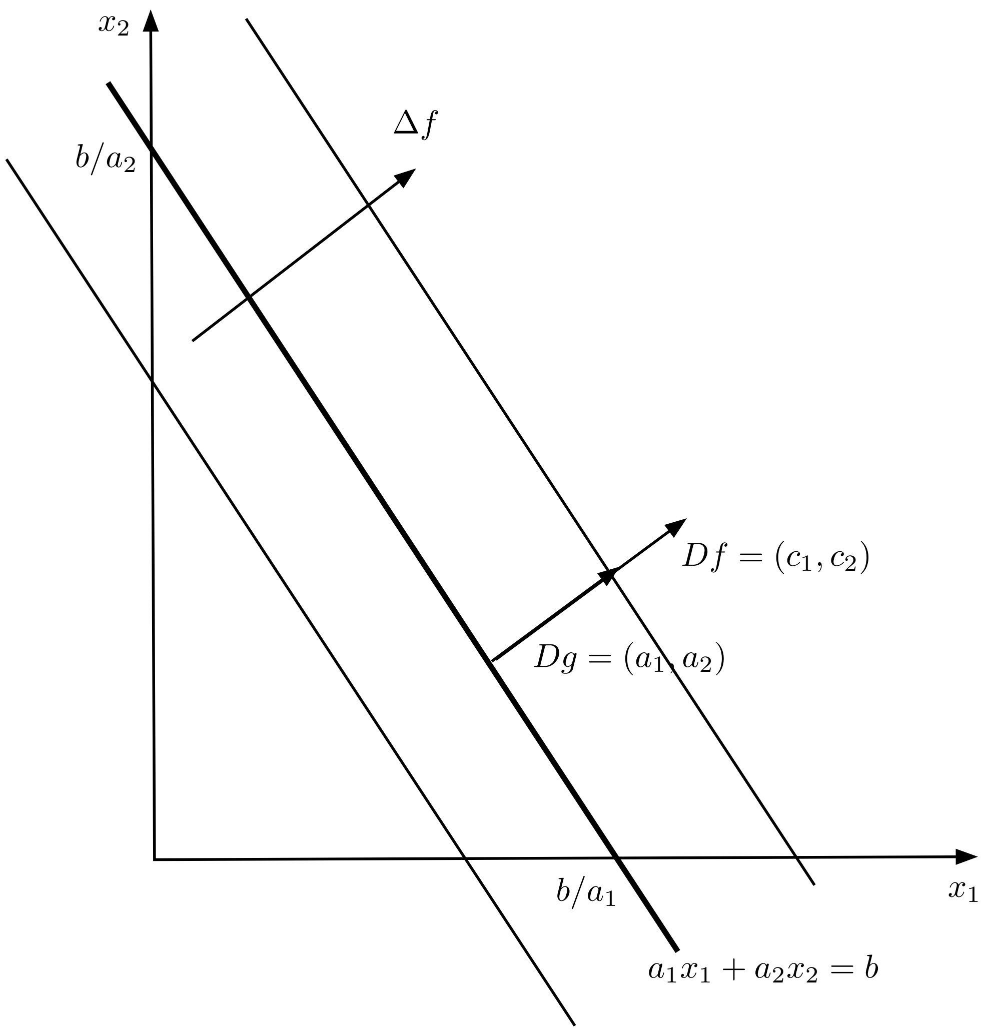
An example violating the constraint qualification (2)



An example satisfying the constraint qualification

- $n = 2, \; m = 1$
- $\max_{x_1,x_2} c_1 x_1 + c_2 x_2$ s.t. $a_1 x_1 + a_2 x_2 = b, \; c_i, a_i > 0$
- $L(x_1, x_2, \lambda) = c_1 x_1 + c_2 x_2 + \lambda(b - a_1 x_1 - a_2 x_2)$
- $\frac{\partial L}{\partial x_1} = c_1 - \lambda a_1 = 0$
- $\frac{\partial L}{\partial x_2} = c_2 - \lambda a_2 = 0$
- $\frac{\partial L}{\partial \lambda} = b - a_1 x_1 - a_2 x_2 = 0$
- from the first two conditions we obtain $\frac{c_1}{a_1} = \lambda = \frac{c_2}{a_2}$
- Then, $a_2 = \frac{c_2 a_1}{c_1}$ and restriction is $x_2 = \frac{b}{a_2} - \frac{c_1}{c_2} x_1$
- $Jf = (c_1, c_2)$.
- $Jg = (a_1, a_2) = (a_1, \frac{a_1 c_2}{c_1}) = \frac{a_1}{c_1}(c_1, c_2) = \frac{a_1}{c_1} Jf$.
- Graphical representation:

An example satisfying the constraint qualification (2)



Lagrange method

SOnC

- Hessian evaluated at $(\mathbf{x}^*, \lambda^*)$ of lagrangian function negative semidefinite:
- $$H = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \cdots & \frac{\partial^2 L}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n^2} x_1 & \frac{\partial^2 L}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}$$

Necessary and sufficient conditions

- $\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^*) = \lambda^* \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}^*)$
- $\mathbf{b} = \mathbf{g}(\mathbf{x}^*)$
- Hessian negative definite.

On lagrange multipliers

- *Proposition*: Lagrange multipliers evaluated at the solution provide a measure of the sensibility of the optimal value of the objective function $f(\mathbf{x}^*)$ to variations in the constants $\mathbf{b}$ of each restriction, namely
  $\lambda^* = \frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{b}}$, or $\lambda_i^* = \frac{\partial f(\mathbf{x}^*)}{\partial b_i}$, $i = 1, 2, \ldots, m$.
- Proof
  - FOCs: $m + n$ equations and $2m + n$ variables $(\mathbf{b}, \lambda, \mathbf{x})$.
  - Implicit function theorem: solve system of $m + n$ equations as function of constants $\mathbf{b}$, i.e.
    $\lambda = \lambda(\mathbf{b});\ \mathbf{x} = \mathbf{x}(\mathbf{b})$.
  - Rewrite lagrangian function as:
    $L(\mathbf{b}) = f(\mathbf{x}(\mathbf{b})) + \lambda(\mathbf{b})[\mathbf{b} - \mathbf{g}(\mathbf{x}(\mathbf{b}))]$
  - Differentiate $L(\mathbf{b})$ wrt $\mathbf{b}$
  - $\frac{\partial L}{\partial \mathbf{b}} = \left(\frac{\partial f}{\partial \mathbf{x}} - \lambda \frac{\partial g}{\partial \mathbf{x}}\right)\left(\frac{\partial \mathbf{x}}{\partial \mathbf{b}}\right) + (\mathbf{b} - \mathbf{g}(\mathbf{x}))' \frac{\partial \lambda'}{\partial \mathbf{b}} + \lambda$

On lagrange multipliers (2)

- Proof (cont'd)
  - Evaluated at $(\mathbf{x}^*, \lambda^*)$, first two terms $= 0$ from FOCs.
  - Thus, $\frac{\partial L}{\partial \mathbf{b}} = \lambda$
  - Also, evaluated at $(\mathbf{x}^*, \lambda^*)$, $L(\mathbf{x}^*, \lambda^*) = f(\mathbf{x}^*)$.
  - Thus, $\frac{\partial L}{\partial \mathbf{b}}(\mathbf{x}^*, \lambda^*) = \frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{b}}$
  - Hence, $\frac{\partial L}{\partial \mathbf{b}}(\mathbf{x}^*, \lambda^*) = \frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{b}} = \lambda^*$.
- Economic interpretation: objective function (profits, costs, revenues, ...) and restrictions (inputs, ...). Then,
  - Lagrange multipliers measure sensibility of say cost, to variations in a quantity (of inputs).
  - Thus, lagrange multipliers represent a price, often referred to as “shadow price" of each input.

Lagrange's Theorem - Proof

Theorem (consider $\mathbf{R}^2$)

- Let $f(x_1, x_2)$ and $g(x_1, x_2)$ be continuously differentiable.
- Let $x^* \equiv (x_1^*, x_2^*)$ be an interior point in the domain of f.
- Let x^* be a local extreme point for $f(x_1, x_2)$ subject to $g(x_1, x_2) = c$.
- Suppose that $\frac{\partial g}{\partial x_1}(x^*)$ and $\frac{\partial g}{\partial x_2}(x^*)$ are not both zero.
- Then, $\exists \lambda \in \mathbf{R}$ such that the lagrangean function

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda(g(x_1, x_2) - c)$$

has a stationary point at x^*.

Lagrange's Theorem - Proof (2)

Proof

- Suppose that $\frac{\partial g}{\partial x_2}(x^*) \neq 0$.
- By the implicit function theorem, the equation $g(x_1, x_2) = c$ defines x_2 as a differentiable function of x_1, $x_2 = h(x_1)$ in a neighborhood of x^*. Also, $h'(x_1) = -(\frac{\partial g}{\partial x_1})/(\frac{\partial g}{\partial x_2})$.
- Replacing $g(x_1, x_2) = c$ by $x_2 = h(x_1)$, the problem

$$\max_{x_1, x_2} f(x_1, x_2) \text{ s.t. } g(x_1, x_2) = c \tag{1}$$

becomes

$$\max_{x_1} f(x_1, h(x_1))$$

Lagrange's Theorem - Proof (3)

Proof (cont'd)

- A neccessary condition for a maximum of this free-maximization problem is

$$\frac{df(x_1, h(x_1))}{x_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} h'(x_1) = 0$$

  That is

$$\frac{\partial f}{\partial x_1}(x^*) - \frac{\partial f}{\partial x_2}(x^*) \frac{\frac{\partial g}{\partial x_1}(x^*)}{\frac{\partial g}{\partial x_2}(x^*)} = 0 \tag{2}$$

- Then, the expression (2) is precisely the necessary condition for x^* to be a local (interior) extreme point of the problem (1).

FOC of problem (1)

- $\frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} = 0$
- $\frac{\partial f}{\partial x_2} - \lambda \frac{\partial g}{\partial x_2} = 0$
- That is,

$$\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} = \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}}$$

- that we can rewrite as,

$$\frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2} \frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} = 0$$

that is precisely expression (2).

Lagrange's Theorem - Proof (4)

Proof (cont'd)

- Accordingly, we conclude that when $\frac{\partial g}{\partial x_2}(x^*) \neq 0$, the Lagrangean function $L(x_1, x_2, \lambda)$ has a stationary point at x^*.
- It remains to prove that x^* is also a stationary point of $L(x_1, x_2, \lambda)$ when $\frac{\partial g}{\partial x_2}(x^*) = 0$. In this case, it follows that $\frac{\partial g}{\partial x_1}(x^*) \neq 0$ and a parallel argument completes the proof of the theorem.

Lagrange's Theorem - Alternative proof

- Let $\mathbf{x}^*$ be a local extreme point for $f(x_1, x_2)$ s.t. $g(x_1, x_2) = c$.
- Represent $g(x_1, x_2) = c$ by the vector valued function $\mathbf{r}(t) = x_1(t)\mathbf{i} + x_2(t)\mathbf{j}$ with $\mathbf{r}(t) \neq \mathbf{0}$
- Define $h(t) = f(x_1(t), x_2(t))$
- Because $f(\mathbf{x}^*)$ is an extreme value of f, $h(t^*) = f(x_1(t^*), x_2(t^*)) = f(\mathbf{x}^*)$ is an extreme value of h.
- It means that $h'(t^*) = 0$
- But $h'(t^*) = \frac{\partial f}{\partial x_1}(\mathbf{x}^*)x_1'(\mathbf{x_1}^*) + \frac{\partial f}{\partial x_2}(\mathbf{x}^*)x_2'(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) \cdot \mathbf{r}'(\mathbf{t}^*)$
- Hence, $\nabla f(\mathbf{x}^*)$ is orthogonal to $\mathbf{r}'(\mathbf{t}^*)$
- Differentiating g we obtain $\frac{\partial g}{\partial x_1}(\mathbf{x}^*)x_1'(\mathbf{x}^*) + \frac{\partial g}{\partial x_2}(\mathbf{x}^*)x_2'(\mathbf{x}^*) = 0$ or $\nabla g(\mathbf{x}^*) \cdot \mathbf{r}'(\mathbf{t}^*) = 0$ and $\nabla g(\mathbf{x}^*)$ is orthogonal to $\mathbf{r}'(\mathbf{t}^*)$
- Thus, $\nabla f(\mathbf{x}^*)$ and $\nabla g(\mathbf{x}^*)$ are parallel. That is, $\exists \lambda^*$ such that $\nabla f(\mathbf{x}^*) = \lambda^* \nabla g(\mathbf{x}^*)$

The Envelope Theorem

Motivation

- Let $F(x, y; \theta)$ be a function with (x, y) as endogenous variables while θ represent some exogenous variable.
- F might be the profit function of a firm producing two outputs (x, y), and θ be the wage rate paid to its workers (determined by labor market conditions).
- Suppose (x^*, y^*) are the profit-maximizing production volumes and $F(x^*, y^*; \theta)$ is the maximum level of profits.
- It should be obviuos that $x^* = x^*(\theta), y^* = y^*(\theta), F(x^*, y^*; \theta) = F^*(\theta)$
- How variation of θ affect F^*? Formally, compute $dF^*/d\theta$.
- Apply the envelope theorem

The Envelope Theorem (2)

The envelope theorem

- Consider $F(x, y; \theta)$, and suppose $x^* = x^*(\theta), y^* = y^*(\theta)$ exist.
- Then,

$$\frac{dF^*}{d\theta} = \frac{\partial F^*}{\partial \theta}$$

Proof

- Compute $dF^*/d\theta$:

$$\frac{dF^*}{d\theta} = \frac{\partial F^*}{\partial x}\frac{dx^*}{d\theta} + \frac{\partial F^*}{\partial y}\frac{dy^*}{d\theta} + \frac{\partial F^*}{\partial \theta}\frac{d\theta}{d\theta} = \frac{\partial F^*}{\partial \theta}$$

where we have used the fact that the FOCs evaluated at the optimal values are zero, i.e.

$$\frac{\partial F}{\partial x}(x^*, y^*; \theta) = 0 = \frac{\partial F}{\partial y}(x^*, y^*; \theta)$$

The Envelope Theorem (3)

Remarks

- The envelope theorem only requires (x^*, y^*) be critical points of F, not optimal.
- An immediate application of the envelope theorem allows us to assess the impact of a variation of the lagrange multiplier on the optimal value of the objective function.

Interpretating the Lagrange multiplier

- Let the Lagrangian function be $F(x, y, \lambda; \theta) = f(x, y) - \lambda(g(x, y) - \theta)$
- Recall that at the optimum, $F^* = f^*$
- Therefore, $\frac{dF^*}{d\theta} = \frac{\partial F^*}{\partial \theta} = \frac{df^*}{d\theta}$
- But we can write $F(x, y, \lambda; \theta) = f(x, y) - \lambda g(x, y) + \lambda\theta$ so that $\frac{dF^*}{d\theta} = \frac{\partial F^*}{\partial \theta} = \lambda^* = \frac{df^*}{d\theta}$