APEC Math Review

Part 6 Optimization

Ling Yao

August, 2020

1. Unconstrained optimization

First-order conditions

We reviewed critical points, global/local max/min in part 4.

Let $F : U \to \mathbb{R}$ be a differentiable defined on a subset U of $\mathbb{R}^n$. If $\mathbf{x}^* \in \mathbb{R}^n$ i a local min or local max of $F(\cdot)$ and if $\mathbf{x}^*$ is a interior point of U, Then

$$\frac{\partial F(\mathbf{x}^*)}{\partial x_n} \quad \textit{for every } n$$

or, in more concise notation

$$\nabla F(\mathbf{x}^*) = \mathbf{0}.$$

Q: Is being a critical point a necessary or sufficient condition for being a local max/min?

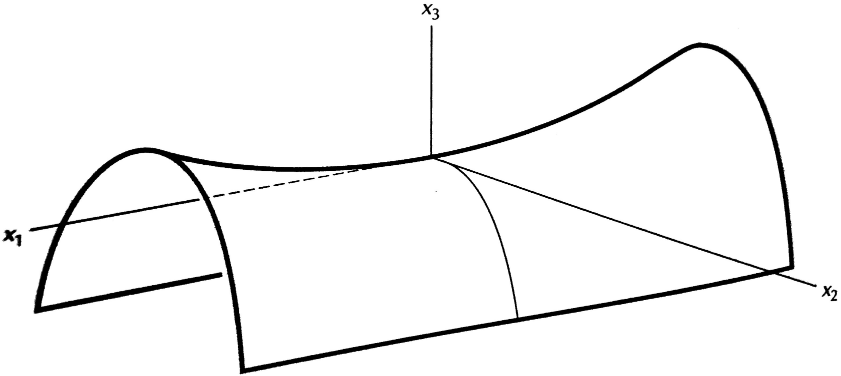
Second-order conditions: sufficient conditions

Let $F : U \to \mathbb{R}$ is C^2 whose domain is open set $U \in \mathbb{R}^n$.
Suppose $\nabla F(\mathbf{x}^*) = \mathbf{0}$

1. If $D^2 f(\mathbf{x}^*)$ is negative (positive) **definite**, then $\mathbf{x}^*$ is strict local max (min).
2. If $D^2 f(\mathbf{x}^*)$ is indefinite, then $\mathbf{x}^*$ is neither a local max or local min.

Q: Why won't a negative semidefinite $D^2 f(\mathbf{x}^*)$ work?

Second-order conditions: sufficient conditions



The graph of the indefinite form $Q_3(x_1, x_2) = x_1^2 - x_2^2$.

Source: Simon & Blume page 378

Second-order conditions: sufficient conditions

Proof: For an arbitrary vector $\mathbf{z} \in \mathbb{R}^n$ and scalar t, a Taylor's expansion of the function $g(t) = f(\mathbf{x}^* + t\mathbf{z})$ around $t = 0$ gives

$$\begin{aligned} f(\mathbf{x}^* + t\mathbf{z}) &= f(\mathbf{x}^*) + t\nabla f(\mathbf{x}^*) \cdot \mathbf{z} + \frac{1}{2}t^2\mathbf{z}^T D^2 f(\mathbf{x}^*)\mathbf{z} + R_2 \\ &= f(\mathbf{x}^*) + \frac{1}{2}t^2\mathbf{z}^T D^2 f(\mathbf{x}^*)\mathbf{z} + R_2 \end{aligned}$$

The remainder is small if t is small, so

$$\mathbf{z}^T D^2 f(\mathbf{x}^*)\mathbf{z} \leq 0$$

Similarly if $\mathbf{z}^T D^2 f(\mathbf{x}^*)\mathbf{z} \leq 0$ for any $\mathbf{z} \neq \mathbf{0}$, then $f(\mathbf{x}^* + t\mathbf{z}) - f(\mathbf{x}^*) < 0$ for small $t > 0$, and so $\mathbf{x}^*$ is a local maximizer.

See Simon & Blume page 838 for details about the remainder.

Second-order conditions: necessary conditions

Let $F : U \to \mathbb{R}$ is C^2 whose domain is open set $U \in \mathbb{R}^n$. Suppose If $\mathbf{x}^* \in \mathbb{R}^n$ is a local max(min) of F. Then, $\nabla F(\mathbf{x}^*) = \mathbf{0}$ and the (symmetric) $n \times n$ matrix $D^2 f(\mathbf{x}^*)$ is negative (positive) **semidefinite.**

Global max and min

Any point $\mathbf{x}^*$ of a concave (convex) function $f(\cdot)$ satisfying $\nabla f(\mathbf{x}^*) = \mathbf{0}$ is a global max (min) of $f(\cdot)$.

Prove it as an exercise.

Application: Profit maximization

Suppose a firm uses n inputs to produce a single product. $\mathbf{x} \in \mathbb{R}^n$ represents an input bundle. $y = Q(\mathbf{x})$ is the production function. p is the selling price of the product and $\mathbf{w}$ is the cost of inputs. The firm's profit function is

$$\pi(\mathbf{x}) = pQ(\mathbf{x}) - \mathbf{w}\mathbf{x}$$

First order conditions

$$\frac{\partial \pi}{\partial x_i}(\mathbf{x}^*) = 0$$

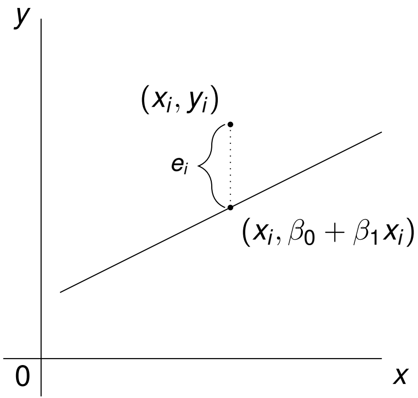
What does this imply? What is the second order necessary conditions? What does it imply?

Application: OLS

Suppose we want to estimate the following single variable linear model with N observations

$$y = \beta_0 + \beta_1 x + e$$

Our goal is to minimize the sum of the squared estimation error. Derive the estimator of β_0 and β_1.



2. Optimization s.t. equality constraints

Lagrange's method: two variables, one constraint

Let f and h be C^1 function of two variables. Suppose that $\mathbf{x}^* = (x_1^*, x_2^*)$ is a solution of the problem

$$\textit{maximize} \quad f(x_1, x_2)$$

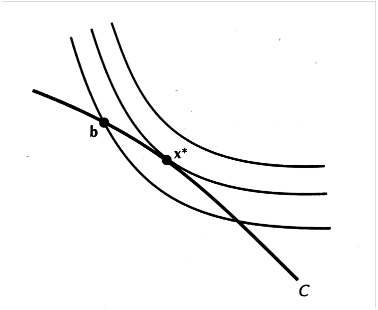
$$s.t. \quad h(x_1, x_2) = c$$

Suppose further that (x_1^*, x_2^*) is not a critical point of h. Then there is a real number μ^* such that (x_1^*, x_2^*, μ^*) is a critical point of the Lagrangian function

$$L(x_1, x_2, \mu) \equiv f(x_1, x_2) - \mu[h(x_1, x_2) - c]$$

In other words, at (x_1^*, x_2^*, μ^*)

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial \mu} = 0$$



Source: Simon & Blume page 414

At $\mathbf{x}^*$

$$-\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)} = -\frac{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}$$

Lagrange's method: intuition

Let

$$\frac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \frac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)} = \mu$$

Then we have two equations

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) - \mu\frac{\partial h}{\partial x_1}(\mathbf{x}^*) = 0$$

$$\frac{\partial f}{\partial x_2}(\mathbf{x}^*) - \mu\frac{\partial h}{\partial x_2}(\mathbf{x}^*) = 0$$

In gradient notation

$$\nabla f(\mathbf{x}^*) = \mu^* \nabla h(\mathbf{x}^*)$$

Q: why (x_1^*, x_2^*) cannot be a critical point of h? When will this be satisfied?

Lagrange's method: multiple variables and constraints

Let $f, h_1, ..., h_m$ be C^1 functions of n variables. Consider the problem of maximizing (or minimizing) $f(\mathbf{x})$ on the constraint set

$$C_{\mathbf{h}} = \{\mathbf{x} = (x_1, ..., x_n) : h_1(\mathbf{x} = a_1, ..., h_m(\mathbf{x} = a_m)\}$$

Suppose that $\mathbf{x}^* \in C_{\mathbf{h}}$ and it is a (local) max or min of f on $C_{\mathbf{h}}$. Suppose further that $\mathbf{x}^*$ is not the critical point of $\mathbf{h} = (h_1, ..., h_m)$ (i.e.the rank of $D\mathbf{h}(\mathbf{x}^*)$ is $< m$). Then there exists real numbers $\mu_1^*, ...\mu_m^*$ such that $(x_1^*, ..., x_n^*, \mu_1^*, ...\mu_m^*)$ is a critical point of the Lagrangian function

$$L(\mathbf{x}^*, \boldsymbol{\mu}^*) \equiv f(\mathbf{x}) - \mu_1[h(\mathbf{x}) - a_1] - \cdots - \mu_m[h(\mathbf{x}) - a_m]$$

In other words, at (x_1^*, x_2^*, μ^*)

$$\frac{\partial L}{\partial x_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, ..., \quad \frac{\partial L}{\partial x_n}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0$$

$$\frac{\partial L}{\partial \mu_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, ..., \quad \frac{\partial L}{\partial \mu_m}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0$$

Exercise: Lagrange's method

(Simon & Blume exercise 18.7)
Maximize $f(x, y, z) = yz + xz$ subject to $y^2 + z^2 = 1$ and $xz = 3$.

Second-order conditions: two variables, one constraint

- With the first order conditions, we can find out the critical points for the Lagrangian function $L(\mathbf{x}, \mu)$.
- We need to know whether they are max or min.
- Are the second order conditions about the Hessian of $L(\mathbf{x}, \mu)$?
- Turns out it is more stringent than we need, because we can exploit the interdependence between the **x**s imposed by the constraint.

Second-order conditions: two variables, one constraint

To know that we have a maximum, all we really need is that the second differential of the objective function at the critical point is decreasing **along the constraint.**

By the implicit function theorem,

$$\frac{dx_2}{dx_1} = -\frac{\partial h/\partial x_1}{\partial h/\partial x_2}$$

Let $y = f(x_1, x_2(x_1))$ be the value of objective function subject to the constraint. By the chain rule,

$$\frac{dy}{dx_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dx_1} = \frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\frac{\partial h/\partial x_1}{\partial h/\partial x_2}$$

Second-order conditions: two variables, one constraint

The second order sufficient condition requires that

$$\frac{d^2y}{dx_1^2} < 0$$

It can be shown that

$$\frac{d^2y}{dx_1^2} = \frac{-1}{(\partial h/\partial x_2)^2}\bar{D}$$

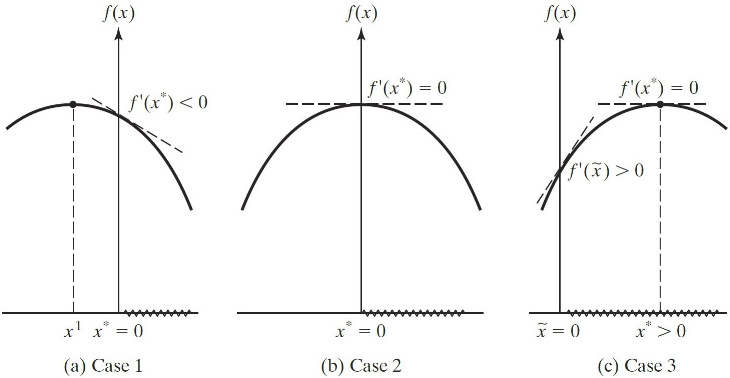
where $\bar{D}$ is the determinant of a **boarded Hessian** of L

$$\begin{pmatrix} 0 & \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} \\ \frac{\partial h}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial h}{\partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \end{pmatrix}$$

3.Optimization s.t. inequality constraints

A simple example

$$\max_{x} f(x) \quad s.t. x \geq 0$$

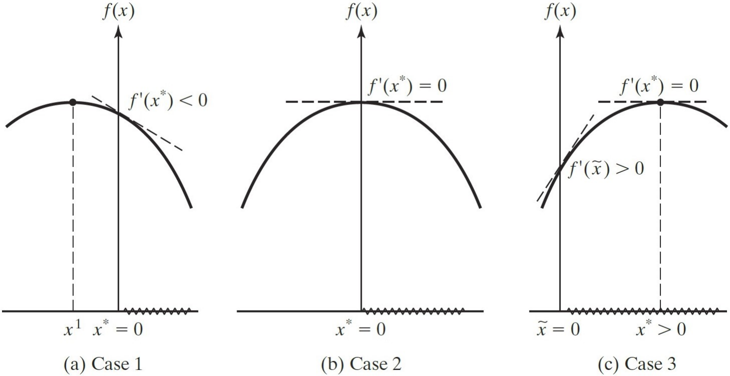


Source: Figure A2.9 in Jehle & Reny (2011)

In any case, $x^*[f'(x^*)] = 0$

A simple example

Question: In which of these cases is the constraint binding?



Source: Figure A2.9 in Jehle & Reny (2011)

Maximization

Necessary conditions for optimal of real-valued functions subject to non-negativity constraints

Let $f(\mathbf{x})$ be continuously differentiable.If $\mathbf{x}^*$ maximizes $f(\mathbf{x})$ subject to $\mathbf{x} \geq 0$, then $\mathbf{x}^*$ satisfies

1. $\frac{\partial f(\mathbf{x})}{\partial x_i} \leq 0,\ i = 1, ..., n$
2. $x_i^*[\frac{\partial f(\mathbf{x})}{\partial x_i}] = 0,\ i = 1, ..., n$
3. $x_i^* \geq 0,\ i = 1, ..., n$

Minimization

Necessary conditions for optimal of real-valued functions subject to non-negativity constrains

Let $f(\mathbf{x})$ be continuously differentiable.If $\mathbf{x}^*$ minimizes $f(\mathbf{x})$ subject to $\mathbf{x} \geq 0$, then $\mathbf{x}^*$ satisfies

1. $\frac{\partial f(\mathbf{x})}{\partial x_i} \geq 0,\ i = 1, ..., n$
2. $x_i^* [\frac{\partial f(\mathbf{x})}{\partial x_i}] = 0,\ i = 1, ..., n$
3. $x_i^* \geq 0,\ i = 1, ..., n$

Non-negativity constraints + other inequality constraints

$$\max_{\mathbf{x}\in\mathbb{R}^n_+} f(\mathbf{x}) \quad s.t.\, \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \quad \mathbf{x} \geq 0$$

$$\tilde{L} = f(\mathbf{x}) - \lambda_1[g_1(\mathbf{x}) - b_1] - ... - \lambda_k[g_k(\mathbf{x}) - b_k]$$

F.O.C. in terms of the Kuhn-Tucker Lagrangian

$\frac{\partial \tilde{L}}{\partial x_i^*} \leq 0,$
$x_i^* \frac{\partial \tilde{L}}{\partial x_i^*} = 0,$
$x_i^* \geq 0$
for $i = 1, ..., n$

$\frac{\partial \tilde{L}}{\partial \lambda_j^*} \geq 0,$
$\lambda_j^* \frac{\partial \tilde{L}}{\partial \lambda_j^*} = 0,$
$\lambda_j^* \geq 0$
for $j = 1, ..., k$

For minimization, simply substitute $f(\mathbf{x})$ with $-f(\mathbf{x})$.

Complementary slackness

$\lambda_j^* \frac{\partial \tilde{L}}{\partial \lambda_j^*} = 0$ implies that at least one of λ_j^* and $\frac{\partial \tilde{L}}{\partial \lambda_j^*}$ must be zero.

- If the constraint is not binding ($\frac{\partial \tilde{L}}{\partial \lambda_j^*} \equiv b_j - g_j(\mathbf{x}) > 0$), then λ_j^* must be zero.
- If $\lambda_j^* > 0$, then the constraint must be binding ($b_j = g_j(\mathbf{x})$).

Application: Corner solution

$$\max_{x_1, x_2 \in \mathbb{R}^n_+} U(x_1, x_2) \quad s.t. \, p_1 x_1 + p_2 x_2 \leq y,$$

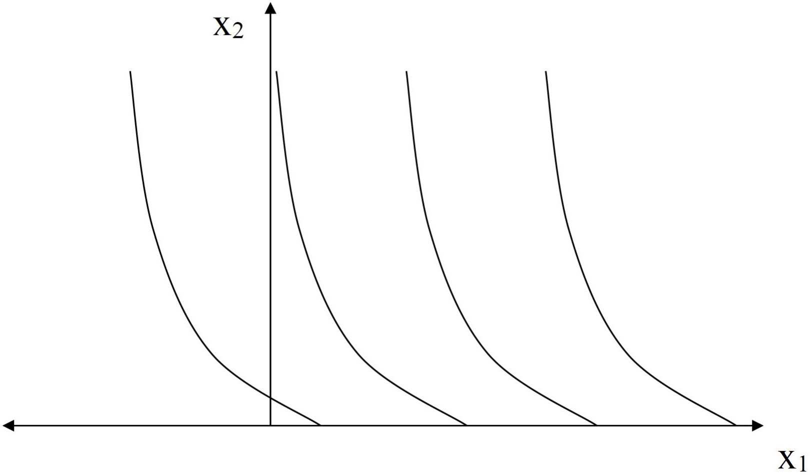
$$L = U(x_1, x_2) - \lambda(p_1 x_1 + p_2 x_2 - y)$$

F.O.C

$$\begin{aligned}
\frac{\partial \tilde{L}}{\partial x_1^*} &= MU_1^* - \lambda^* p_1 \leq 0 \qquad & x_1^* \frac{\partial \tilde{L}}{\partial x_1^*} &= 0 \qquad & x_1^* &\geq 0 \\
\frac{\partial \tilde{L}}{\partial x_2^*} &= MU_2^* - \lambda^* p_2 \leq 0 \qquad & x_2^* \frac{\partial \tilde{L}}{\partial x_2^*} &= 0 \qquad & x_2^* &\geq 0 \\
\frac{\partial \tilde{L}}{\partial \lambda^*} &= -p_1 x_1^* - p_2 x_2^* + y \geq 0 \qquad & \lambda^* \frac{\partial \tilde{L}}{\partial \lambda^*} &= 0 \qquad & \lambda^* &\geq 0
\end{aligned}$$

If $x_i^* = 0$ then MU_i^* can deviate from $\lambda^* p_i$

Application: Corner solution



Indifference curves of a quasilinear preference

Exercise: Kuhn-Tucker conditions

(Simon & Blume Example 18.13)
Solve for the problem of maximizing $f(x, y) = x^2 + x + 4y^2$ subject to the inequality constraints

$$2x + 2y \leq 1, \quad x \geq 0, \quad y \geq 0$$

3. Comparative statics and the envelope theorem

The meaning of the multiplier

Consider a two variables, one equality constraint problem:

$$max \quad f(x, y) s.t. \quad h(x, y) = a$$

For a fixed value of the parameter a, let $(x^*(a), y^*(a))$ be the solution of the problem with corresponding multiplier $\mu^*(a)$. Suppose that x^*, y^* and μ^* are C^1 functions of a, then

$$\mu^*(a) = \frac{d}{da} f(x^*(a), y^*(a)).$$

In the case of multiple variables and multiple equality constraints,

$$\mu_j^*(a_1, ..., a_m) = \frac{\partial}{\partial a_j} f(x_1^*(a_1, ..., a_m), ..., x_n^*(a_1, ..., a_m))$$

for each $j = 1, ..., m$.

Application

What are the interpretations of the Lagrange multiplier in the following problems:

- Utility maximization subject to budget constraint
- Profit maximization subject to input availability constraint

The Envelope Theorem: unconstrained

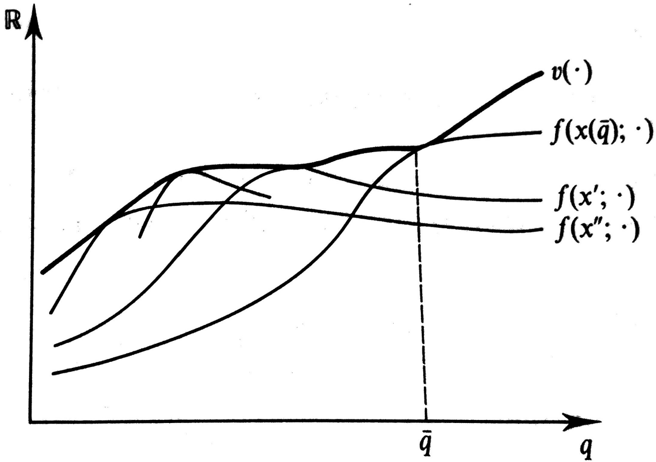
Let $f(\mathbf{x}; a)$ be a C^1 function of $\mathbf{x} \in \mathbb{R}^n$ and the scalar a. For each choice of the parameter a, consider the unconstrained optimization problem

$$maximize \quad f(\mathbf{x}; a) \quad w.r.t. x$$

Let $\mathbf{x}^*(a)$ be a solution of this problem. Suppose that $\mathbf{x}^*(a)$ is a C^1 function of a. Then,

$$\frac{d}{da} f(\mathbf{x}^*(a); a) = \frac{\partial}{\partial a} f(\mathbf{x}^*(a); a)$$

The Envelope Theorem: intuition



Source: M.W.G. p965

Example

$$\textit{maximize} \quad f(x,a) = -x^2 + 2ax + 4a^2$$

F.O.C.

$$f'(x) = -2x + 2a = 0$$

$$x^* = a$$

Plugging this back into $f(x,a)$ we get a single variable function

$$f(x^*(a);a) = f(a,a) = -a^2 + 2a \cdot a + 4a^2 = 5a^2$$

So

$$\frac{df^*}{da} = 10a$$

This is equal to the partial derivative of the original function at the optimum

$$\frac{\partial f(x^*(a),a)}{\partial a} = 2x + 8a = 10a.$$

The Envelope Theorem: constrained

Let $f, h_1, ..., h_m : \mathbb{R}^n \times \mathbb{R}^1 \to \mathbb{R}^1$ be C^1 functions. Let $\mathbf{x}^*(a) = (x_1^*(a), ..., x_n^*(a))$ denote the solution of the problem of maximizing $f(\mathbf{x}; a)$ with respect to $\mathbf{x}$ on the constraint set

$$h_1(\mathbf{x}; a) = 0, ..., h_m(\mathbf{x}; a) = 0$$

for any fixed choice of parameter a. Suppose that $\mathbf{x}^*(a)$ and the Lagrange multipliers $\mu_1(a), ..., \mu_m(a)$ are C^1 functions of a. Then

$$\frac{d}{da} f(\mathbf{x}^*(a); a) = \frac{\partial L}{\partial a}(\mathbf{x}^*, \mu(a); a)$$

Exercise

Verify that the interpretation of the Lagrange multiplier is a special case of the envelope function theorem using the example problem:

$$\max_{x_1, x_2} f(x_1, x_2) = x_1 x_2 \quad s.t \quad 2x_1 + 4x_2 = a$$

Exercise

In the utility maximizing problem, the Roy's identity says that the consumer's demand for good *i* is the ratio of the partial derivatives of the maximized utility with respect to good *i*'s price and income with a minus sign, i.e.

$$-\frac{\partial u^*(\mathbf{p}, y)/\partial p_i}{\partial u^*(\mathbf{p}, y)/\partial y} = x_i^* = \mathbf{x}_i(\mathbf{p}, y)$$

Prove it using the Envelope theorem.