

Lecture 3: Maximization and Comparative Statics

3.1 Maximization with many variables

3.1.1 Theorem (Necessary First Order Conditions) *If U is an open subset of a normed space, and $x^* \in U$ is a local extremum of f , and f has directional derivatives at x^* , then for any nonzero v , the directional derivative satisfies $D_v f(x^*) = 0$. In particular, if f is differentiable at x^* , then $Df(x^*) = 0$.*

Proof using one-dimensional case: Since x^* is an interior point of U , there is an $\varepsilon > 0$ such that $x^* + \lambda v \in U$ for any $\lambda \in (-\varepsilon, \varepsilon)$ and any $v \in \mathbf{R}^n$ with $\|v\| = 1$. Set $g_v(\lambda) = f(x^* + \lambda v)$. Then g_v has an extremum at $\lambda = 0$. Therefore $g'_v(0) = 0$. By the chain rule, $g'_v(\lambda) = Df(x^* + \lambda v)(v)$. Thus we see that $Df(x^*)(v) = 0$ for every v , so $Df(x^*) = 0$. ■

3.1.2 Theorem (Necessary second order conditions) *Let f be a continuously differentiable real-valued function on an open subset U of \mathbf{R}^n and assume that f is twice differentiable at x^* , and define the quadratic form $Q(v) = D^2 f(x^*)(v, v)$. If x^* is a local maximizer, then Q is negative semidefinite. If x^* is a local minimizer, then Q is positive semidefinite.*

Proof using the chain rule: As in the proof of Theorem 3.1.1, define $g(\lambda) = f(x^* + \lambda v)$. It achieves a maximum at $\lambda = 0$, so the second order condition is $g''(0) \leq 0$. So by the chain rule, using $Df(x^*) = 0$,

$$g''(0) = D^2 f(x^*)(v, v) \leq 0.$$

That is, Q is negative semidefinite. ■

3.1.3 Theorem (Sufficient second order conditions) *Let f be a continuously differentiable real-valued function on an open subset U of \mathbf{R}^n . Let x^* belong to U and assume that $Df(x^*) = 0$ and that f is twice differentiable at x^* .*

If the Hessian matrix $f''(x^)$ is positive definite, then x^* is a strict local minimizer of f .*

If the Hessian matrix $f''(x^)$ is negative definite, then x^* is a strict local maximizer of f .*

If the Hessian is nonsingular but indefinite, then x^ is neither a local maximum, nor a local minimum.*

Use $D^2 f(x)$.

See my extended on-lines notes [2], especially Section 3.14.

- Strong second order conditions and the Implicit Function Theorem.

3.2 Comparative statics of first order conditions

Start with a function $f: X \times P \rightarrow \mathbf{R}$ where $X \subset \mathbf{R}^n$ and $P \subset \mathbf{R}^m$. For each $p \in P$ let $x^*(p)$ be the interior maximizer of $f(\cdot; p)$. The the first order conditions

$$\begin{aligned} \frac{\partial}{\partial x_1} f(x_1^*(p_1, \dots, p_m), \dots, x_n^*(p_1, \dots, p_m); p_1, \dots, p_m) &= 0 \\ &\vdots \\ \frac{\partial}{\partial x_i} f(x_1^*(p_1, \dots, p_m), \dots, x_n^*(p_1, \dots, p_m); p_1, \dots, p_m) &= 0 \\ &\vdots \\ \frac{\partial}{\partial x_n} f(x_1^*(p_1, \dots, p_m), \dots, x_n^*(p_1, \dots, p_m); p_1, \dots, p_m) &= 0 \end{aligned}$$

hold for each p . As such the left-hand side is a constant (zero) function of p , and so its partial derivatives are all zero. Differentiating the left-hand side of the i^{th} first order condition with respect to p_k thus gives

$$\sum_{j=1}^n \left[\frac{\partial^2 f(x^*(p); p)}{\partial x_i \partial x_j} \frac{\partial x_j^*(p)}{\partial p_k} \right] + \frac{\partial^2 f(x^*(p); p)}{\partial x_i \partial p_k} = 0$$

for all $i = 1, \dots, n$ and $k = 1, \dots, m$.

In matrix terms this becomes

$$\begin{bmatrix} \frac{\partial^2 f(x^*(p); p)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x^*(p); p)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_n^2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^*(p)}{\partial p_1} & \dots & \frac{\partial x_1^*(p)}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial x_n^*(p)}{\partial p_1} & \dots & \frac{\partial x_n^*(p)}{\partial p_m} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 f(x^*(p); p)}{\partial x_1 \partial p_1} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_1 \partial p_m} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x^*(p); p)}{\partial x_n \partial p_1} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_n \partial p_m} \end{bmatrix}$$

Now if the leftmost matrix has an inverse, then we may write

$$\begin{bmatrix} \frac{\partial x_1^*(p)}{\partial p_1} & \dots & \frac{\partial x_1^*(p)}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial x_n^*(p)}{\partial p_1} & \dots & \frac{\partial x_n^*(p)}{\partial p_m} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 f(x^*(p); p)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x^*(p); p)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_n^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial^2 f(x^*(p); p)}{\partial x_1 \partial p_1} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_1 \partial p_m} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x^*(p); p)}{\partial x_n \partial p_1} & \dots & \frac{\partial^2 f(x^*(p); p)}{\partial x_n \partial p_m} \end{bmatrix}$$

to solve for all the comparative statics results

$$\frac{\partial x^*}{\partial p_k}$$

When does this inverse exist? By the second order conditions for a maximum the matrix

$$\left[\frac{\partial^2 f(x^*; p)}{\partial x_i \partial x_j} \right] \text{ must negative semidefinite.}$$

This means that it is invertible precisely when it is negative definite. Thus

We can solve for all the comparative statics results whenever the strong second order condition (definiteness) holds.

3.3 The envelope theorem

There is one more incredibly useful theorem, called the Envelope Theorem. I'll start by explaining why it's called the envelope theorem.

Given a one-dimensional parameterized family of curves,

$$f_\alpha: [0, 1] \rightarrow \mathbf{R} \quad \text{where } \alpha \text{ runs over some interval } I,$$

a curve

$$h: [0, 1] \rightarrow \mathbf{R}$$

is the **envelope** of the family if

- i. each point on the curve h is tangent to one of the curves f_α and
- ii. each curve f_α is tangent to h .

This is the definition from Apostol [1, p. 342]. That is, for each α , there is some t and also for each t , there is some α , satisfying $f_\alpha(t) = h(t)$ and $f'_\alpha(t) = h'(t)$. If the correspondence between curves and points on the envelope is one-to-one, then we may regard h as a function of α . However, once we regard h as a function of α rather than t , the tangency condition has to be rewritten. This observation is the celebrated “Wong–Viner theorem.”¹

Now let $f: X \times P \rightarrow \mathbf{R}$, where X and P are real intervals, and consider the problem

$$\underset{x \in X}{\text{maximize}} f(x, p).$$

We may call x the *decision variable* and p the *parameter*, or we may call x the *control* and p the *state*, or we may say that x is *endogenous* and p is *exogenous*. The function f is called the *objective function*.

Let $x^*(p)$ be an interior solution to this maximization problem. Note that it depends on the parameter p . Define

$$V(p) = f(x^*(p), p)$$

for each p . The function V is called the **optimal value function**.

For fixed x , the graph of the function $\varphi_x: P \rightarrow \mathbf{R}$ via

$$\varphi_x(p) = f(x, p)$$

defines a curve (or in higher dimensions of P , a surface).

The value function $V(p)$ satisfies

$$V(p) = f(x^*(p), p) = \max_x \varphi_x(p).$$

¹ According to Samuelson [4, p. 34], Jacob Viner asked his draftsman, one Mr. Wong, to draw the long run cost curve passing through the minimum of each short run cost curve, and tangent to it. Mr. Wong argued that this was impossible, and that the correct interpretation was that the long run curve was the envelope of the short run curves. See also Viner [6].

The Envelope Theorem states that under appropriate conditions, the graph of the value function V will be the envelope of the family of curves

$$\{\varphi_x : x \in \text{range } x^*\}.$$

Envelope theorems in maximization theory are concerned with the tangency conditions this entails.

To get a picture of this result, imagine a plot of the graph of f . It is the surface $z = f(x, p)$ in (x, p, z) -space. Orient the graph so that the x -axis is perpendicular to the page and the p -axis runs horizontally across the page, and the z -axis is vertical. The high points of the surface (minus perspective effects) determine the graph of the value function V . Here is an example:

3.3.1 Example Let

$$f(x, p) = p - (x - p)^2 + 1, \quad 0 \leq x, p \leq 2.$$

See Figure 3.1. Then given p , the maximizing x is given by $x^*(p) = p$, and $V(p) = p + 1$. The side-view of this graph in Figure 3.2 shows that the high points do indeed lie on the line $z = 1 + p$. For each x , the function φ_x is given by

$$\varphi_x(p) = p - (x - p)^2 + 1.$$

The graphs of these functions and of V are shown for selected values of x in Figure 3.3. Note that the graph of V is the envelope of the family of graphs φ_x . Moreover the slope of V is given by

$$V'(p) = \left. \frac{\partial f}{\partial p} \right|_{x=x^*(p)=p} = 1 + 2(x - p) \Big|_{x=p} = 1.$$

This last observation is known as the Envelope Theorem. □

3.3.2 Envelope Theorem v. 1 *Assume that f and x^* are differentiable. Then*

$$V'(p) = D_2 f(x^*(p), p).$$

That is, the derivative of the optimal value function is simply the partial derivative of the objective function, evaluated at the optimal decision.

Proof: By hypothesis f and x^* are differentiable, so V is also differentiable. By the chain rule

$$V'(p) = D_1 f(x^*(p), p) \cdot x^{*'}(p) + D_2 f(x^*(p), p),$$

but $D_1 f(x^*(p), p) = 0$ by the necessary first order conditions. ■

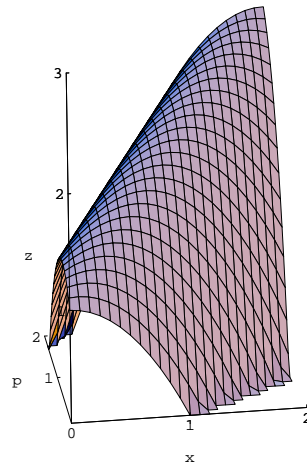


Figure 3.1. Graph of $f(x, p) = p - (x - p)^2 + 1$.

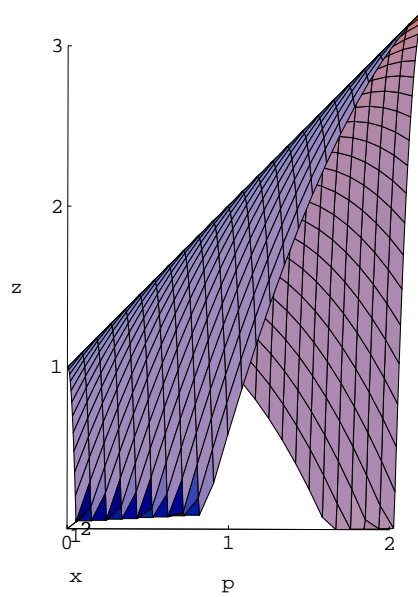


Figure 3.2. Graph of $f(x, p) = p - (x - p)^2 + 1$ viewed from the side.

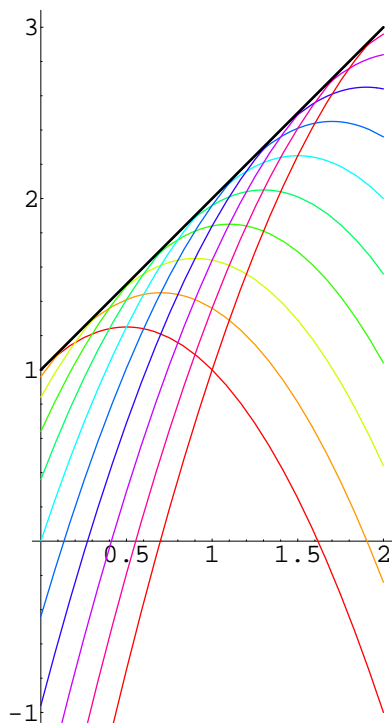


Figure 3.3. Graph of $V(p) = p + 1$ as the envelope of the family $\{\varphi_x(p) : x = 0, .25, \dots, 2\}$, where $\varphi_x(p) = p - (x - p)^2 + 1 = f(x, p)$.

Alternate proof for the one-dimensional case: By hypothesis f and x^* are differentiable, so V is also differentiable. Fix p_0 in the interior of P , and fix $x_0 = x^*(p_0)$. By definition of V , for any p ,

$$V(p) \geq f(x_0, p) \quad \text{and} \quad V(p_0) = f(x_0, p_0).$$

Therefore the function $h(p)$ defined by

$$h(p) = V(p) - f(x_0, p)$$

achieves its minimum at p_0 . The necessary first order condition for this is that

$$V'(p_0) - D_2f(x_0, p_0) = 0.$$

■

Note that the alternative proof generalizes easily to higher dimensional sets P , and it does not use the first-order conditions. In fact, as long as V is differentiable and f is differentiable with respect to p , the argument goes through.

3.4 The Envelope Theorem and the Le Chatelier Principle

Consider a producer that produces output with capital K and labor L according to the production function f .

$$\underset{K,L}{\text{maximize}} pf(K, L) - wL - rK$$

Let $\pi^*(p, w, r)$ be the optimal profit function, and $K^*(p, w, r)$ and $L^*(p, w, r)$ be the **input demand functions**, and $y^*(p, w, r)$ be the **supply function**.

Now suppose K is fixed at $\bar{K} = K^*(\bar{p}, \bar{w}, \bar{r})$. Now the producer wants to maximize

$$\underset{L}{\text{maximize}} pf(\bar{K}, L) - wL - r\bar{K}$$

Let $\hat{L}(p, w, r, \bar{K})$ be the **short-run demand** for labor, $\hat{y}(p, w, r, \bar{K})$ be the **short-run supply**, and $\hat{\pi}$ be the **short-run profit** function. What can we say about

$$\frac{\partial L^*}{\partial p} \text{ vs. } \frac{\partial \hat{L}}{\partial p}, \text{ etc. ?}$$

Fix \bar{w} and \bar{r} . Then

$$\pi^*(p, \bar{w}, \bar{r}) \geq \hat{\pi}(p, \bar{w}, \bar{r}, \bar{K})$$

with equality at $p = \bar{p}$. That is, \bar{p} minimizes the difference, so

$$\frac{\partial \pi^*(\bar{p}, \bar{w}, \bar{r})}{\partial p} - \frac{\partial \hat{\pi}(\bar{p}, \bar{w}, \bar{r}, \bar{K})}{\partial p} = 0,$$

and

$$\frac{\partial^2 \pi^*(\bar{p}, \bar{w}, \bar{r})}{\partial p^2} - \frac{\partial^2 \hat{\pi}(\bar{p}, \bar{w}, \bar{r}, \bar{K})}{\partial p^2} \geq 0.$$

But by the Envelope Theorem,

$$\partial \pi^* / \partial p = \partial (pf(K, L) - wL - rK) / \partial p = f(K^*, L^*) = y^*,$$

so

$$\frac{\partial^2 \pi^*(\bar{p}, \bar{w}, \bar{r})}{\partial p^2} = \frac{\partial y^*(\bar{p}, \bar{w}, \bar{r})}{\partial p},$$

and

$$\frac{\partial^2 \hat{\pi}(\bar{p}, \bar{w}, \bar{r}, \bar{K})}{\partial p^2} = \frac{\partial \hat{y}(\bar{p}, \bar{w}, \bar{r}, \bar{K})}{\partial p}.$$

Thus

$$\frac{\partial y^*(\bar{p}, \bar{w}, \bar{r})}{\partial p} \geq \frac{\partial \hat{y}(\bar{p}, \bar{w}, \bar{r}, \bar{K})}{\partial p}.$$

This sort of result is known as **Le Chatelier's Principle**. (Named for Henry Louis Le Chatelier [3].) Similarly, we can prove

$$\frac{\partial L^*(\bar{p}, \bar{w}, \bar{r})}{\partial w} \leq \frac{\partial \hat{L}(\bar{p}, \bar{w}, \bar{r}, \bar{K})}{\partial w}.$$

3.5 Le Chatelier without the Envelope Theorem

3.5.1 Long Run

$$\text{maximize } R(L, K) - wL - rK$$

FOC

$$\begin{aligned} R_L - w &= 0 \\ R_K - r &= 0 \end{aligned}$$

SOC

$$\begin{bmatrix} R_{LL} & R_{LK} \\ R_{KL} & R_{KK} \end{bmatrix} \text{ is negative semidefinite.}$$

Comparative statics: Differentiate the first order conditions wrt w . Write $L(w)$, $K(w)$.

$$\begin{aligned} R_{LL}L' + R_{LK}K' - 1 &= 0 \\ R_{KL}L' + R_{KK}K' &= 0 \end{aligned}$$

ADD something

$$\begin{bmatrix} R_{LL} & R_{LK} \\ R_{KL} & R_{KK} \end{bmatrix} \begin{bmatrix} L' \\ K' \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \begin{bmatrix} L' \\ K' \end{bmatrix} &= \begin{bmatrix} R_{LL} & R_{LK} \\ R_{KL} & R_{KK} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} L' \\ K' \end{bmatrix} &= \frac{1}{D} \begin{bmatrix} R_{KK} & -R_{LK} \\ -R_{KL} & R_{LL} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

So

$$L' = \frac{R_{KK}}{D} = \frac{R_{KK}}{R_{KK}R_{LL} - R_{KL}^2}$$

By SOC (and existence of inverse) we have

$$R_{KK} < 0, \quad R_{LL} < 0, \quad D > 0.$$

So

$$L'(w) < 0,$$

which is not surprising. Also

$$K'(w) = \frac{-R_{KL}}{D}$$

This sign is harder to figure. It is the opposite of R_{KL} . If w increases, the first order conditions require that the MPL increase. This is accomplished by decreasing L . This in turn changes the MPK by R_{KL} . If $R_{KL} > 0$, then a decrease in L will decrease R_K , so it is now less than r , so K must decrease to raise the MPK up to r .

3.5.2 Short Run

In the short run K is fixed, so the FOC is

$$\begin{aligned} R_{LL}L'_{\text{SR}} - 1 &= 0 \\ L'_{\text{SR}} &= \frac{1}{R_{LL}} \end{aligned}$$

3.5.3 Comparison

How do we compare this to the long run?

In the long run,

$$L' = \frac{R_{KK}}{R_{KK}R_{LL} - R_{KL}^2} = \frac{1}{R_{LL} - \frac{R_{KL}^2}{R_{KK}}} = \frac{1}{R_{LL} + \varepsilon} < 0$$

where $\varepsilon > 0$. Thus

$$0 > L' > L'_{\text{SR}}.$$

That is, the short run response of L to a change in w is greater in magnitude than the long run response.

3.6 Quasiconcave functions

There are weaker notions of convexity that are commonly applied in economic theory.

Have we defined convexity yet?

3.6.1 Definition A function $f: C \rightarrow \mathbf{R}$ on a convex subset C of a vector space is:

- **quasiconcave** if for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$

$$f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}.$$

- **strictly quasiconcave** if for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$

$$f(\lambda x + (1 - \lambda)y) > \min\{f(x), f(y)\}.$$

- **explicitly quasiconcave** or **semistrictly quasiconcave** if it is quasiconcave and in addition, for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$

$$f(x) > f(y) \implies f(\lambda x + (1 - \lambda)y) > \min\{f(x), f(y)\} = f(y).$$

- **quasiconvex** if for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}.$$

- **strictly quasiconvex** if for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$

$$f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\}.$$

- **explicitly quasiconvex** or **semistrictly quasiconvex** if it is quasiconvex and in addition, for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$

$$f(x) < f(y) \implies f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\} = f(y).$$

There are other choices we could have made for the definition based on the next lemma.

3.6.2 Lemma For a function $f: C \rightarrow \mathbf{R}$ on a convex set, the following are equivalent:

1. The function f is quasiconcave.
2. For each $\alpha \in \mathbf{R}$, the strict upper contour set $[f(x) > \alpha]$ is convex, but possibly empty.
3. For each $\alpha \in \mathbf{R}$, the upper contour set $[f(x) \geq \alpha]$ is convex, but possibly empty.

Proof: (1) \implies (2) If f is quasiconcave and x, y in C satisfy $f(x) > \alpha$ and $f(y) > \alpha$, then for each $0 \leq \lambda \leq 1$ we have

$$f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\} > \alpha.$$

(2) \implies (3) Note that

$$[f \geq \alpha] = \bigcap_{n=1}^{\infty} [f > \alpha - \frac{1}{n}],$$

and recall that the intersection of convex sets is convex.

(3) \implies (1) If $[f \geq \alpha]$ is convex for each $\alpha \in \mathbf{R}$, then for $y, z \in C$ put $\alpha = \min\{f(y), f(z)\}$ and note that $f(\lambda y + (1 - \lambda)z)$ belongs to $[f \geq \alpha]$ for each $0 \leq \lambda \leq 1$. \blacksquare

3.6.3 Corollary *A concave function is quasiconcave. A convex function is quasiconvex.*

3.6.4 Lemma *A strictly quasiconcave function is also explicitly quasiconcave. Likewise a strictly quasiconvex function is also explicitly quasiconvex.*

Of course, not every quasiconcave function is concave.

3.6.5 Example (Explicit quasiconcavity) This example sheds some light on the definition of explicit quasiconcavity. Define $f: \mathbf{R} \rightarrow [0, 1]$ by

$$f(x) = \begin{cases} 0 & x = 0 \\ 1 & x \neq 0. \end{cases}$$

If $f(x) > f(y)$, then $f(\lambda x + (1 - \lambda)y) > f(y)$ for every $\lambda \in (0, 1)$ (since $f(x) > f(y)$ implies $y = 0$). But f is not quasiconcave, as $\{x : f(x) \geq 1\}$ is not convex. \square

For a proof of the next fact see my notes for Ec 181.

3.6.6 Fact *Let C be a convex set in \mathbf{R}^m . Let f be a lower semicontinuous quasiconcave function on C that has no local maxima. Then f is explicitly quasiconcave.*

3.6.7 Corollary *Suppose f is concave on a convex neighborhood $C \subset \mathbf{R}^n$ of x^* , and differentiable at x^* . If $f'(x^*) = 0$, then f has a global maximum over C at x^* .*

3.6.8 Theorem (Local maxima of explicitly quasiconcave functions)

Let $f: C \rightarrow \mathbf{R}$ be an explicitly quasiconcave function (C convex). If x^ is a local maximizer of f , then it is a global maximizer of f over C .*

Proof: Let x belong to C and suppose $f(x) > f(x^*)$. Then by the definition of explicit quasiconcavity, for any $1 > \lambda > 0$, $f(\lambda x + (1 - \lambda)x^*) > f(x^*)$. Since $\lambda x + (1 - \lambda)x^* \rightarrow x^*$ as $\lambda \rightarrow 0$ this contradicts the fact that f has a local maximum at x^* . \blacksquare

References

- [1] T. M. Apostol. 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell.
- [2] K. C. Border. 2015. Miscellaneous notes on optimization theory and related topics: Or a dessert course in optimization.
<http://www.its.caltech.edu/~kcborder/Notes/Maximization.pdf>
- [3] M. H. Le Chatelier. 1884. Sur un énoncé général des lois des équilibres chimiques. *Comptes Rendus des Séances de l'Académie des Sciences (Paris)* 99:786–789.
<http://www.bibnum.education.fr/chimie/théorie-chimique/la-loi-de-le-chatelier-1884>
- [4] P. A. Samuelson. 1965. *Foundations of economic analysis*. New York: Athenaeum. Reprint of the 1947 edition published by Harvard University Press.
- [5] E. Silberberg. 1974. A revision of comparative statics methodology in economics, or how to do comparative statics on the back of an envelope. *Journal of Economic Theory* 7(2):159–172. DOI: [10.1016/0022-0531\(74\)90104-5](https://doi.org/10.1016/0022-0531(74)90104-5)
- [6] J. Viner. 1932. Cost curves and supply curves. *Zeitschrift für Nationalökonomie* 3(1):23–46. DOI: [10.1007/BF01316299](https://doi.org/10.1007/BF01316299)