Caltech Division of the Humanities and Social Sciences

Summary Notes on Maximization

KC Border

Fall 2005

v. 2018.10.03::13.35

1 Classical Lagrange Multiplier Theorem

1 Definition *A point x^* is a* ***constrained local maximizer*** *of f subject to the constraints $g_1(x) = \alpha_1$, $g_2(x) = \alpha_2$,..., $g_m(x) = \alpha_m$ in some neighborhood W of x^* if x^* satisfies the constraints and also satisfies $f(x^*) \geqslant f(x)$ for all $x \in W$ that also satisfy the constraints.*

The classical Lagrange Multiplier Theorem on constrained optima for differentiable functions has a simple geometric interpretation, which is easiest to see with a single constraint. Consider a point that maximizes $f(x)$ subject to the equality constraint $g(x) = \alpha$. It should be clear from Figure 1 that at a point where a local maximum occurs, the level curves of f and g must

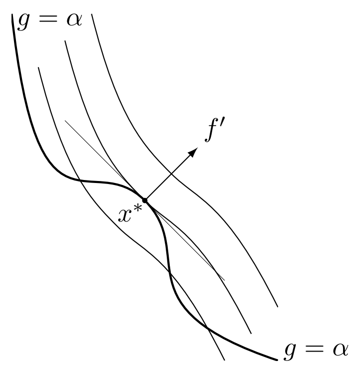

Figure 1. Constrained Maximum with an Equality Constraint.

be tangent. Since the gradient vectors are always perpendicular to the tangent line, they must be colinear. Algebraically, this means that there are coefficients μ^* and λ^* (multipliers, if you will), not both zero, satisfying

$$\mu^* f'(x^*) + \lambda^* g'(x^*) = 0.$$

In general, this is all that can be said. But if the gradient g' is nonzero, then, as we shall see, the multiplier on f' can be taken to be unity, and we get the more familiar condition, $f' + \lambda^* g' = 0$. Note that this does not imply that λ itself is nonzero, since f' may be zero itself. Also note that in general we cannot say anything about the sign of λ^*. That is, there is nothing to tell us if g' points in the same direction as f', or the opposite direction. This changes when we have an inequality constraint. If there is a local maximum of f subject to $g(x) \geqslant \alpha$, then the gradient of g points into $[g > \alpha]$, and the gradient of f points out. See Figure 2. This means that we can take $\mu^*, \lambda^* \geqslant 0$. Even if $[g > \alpha]$ is empty, then $g' = 0$ (why?), so we can take

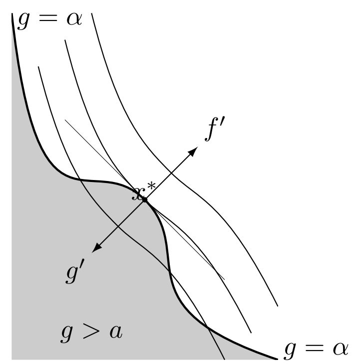

Figure 2. Constrained Maximum with an Inequality Constraint.

$\mu^* = 0$ and $\lambda^* = 1$. That's really all there is to it, so keep these pictures in mind through all the complications needed to express these ideas formally.

The proofs of the Lagrange Multiplier Theorem make use of the Implicit Function Theorem and its corollaries. (See, e.g., my on-line notes [2].) The main result is the Fundamental Lemma on Curves, which says that if x^* satisfies the m constraints $g_1(x), \dots, g_m(x) = 0$, and if v is orthogonal to the gradient of each of the independent constraints at x^*, then there is a differentiable curve $(\hat{x})$ through x^* satisfying the constraints with derivative equal to v at x^*.

2 Lagrange Multiplier Theorem I *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$, and let $f, g_1, \dots, g_m \colon X \to \boldsymbol{R}$ be continuous. Let x^* be an interior constrained local maximizer of f subject to $g(x) = 0$. Suppose f, $g_1, \dots, g_m$ are differentiable at x^*, and that ${g_1}'(x^*), \dots, {g_m}'(x^*)$ are linearly independent.*

Then there exist real numbers $\lambda_1^*, \dots, \lambda_m^*$, such that

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

The next result is provides a different version of the Lagrange Multiplier Theorem that includes the first as a special case. A proof may be found in Carathéodory [3, Theorem 11.1, pp. 175–177].

3 Lagrange Multiplier Theorem II *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$, and let $f, g_1, \dots, g_m \colon X \to \boldsymbol{R}$ be continuous. Let x^* be an interior constrained local maximizer of f subject to $g(x) = 0$. Suppose f, $g_1, \dots, g_m$ are differentiable at x^*.*

Then there exist real numbers $\mu^*, \lambda_1^*, \dots, \lambda_m^*$, not all zero, such that

$$\mu^* f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) = 0.$$

Furthermore, if ${g_1}'(x^*), \dots, {g_m}'(x^*)$, are linearly independent, we may take μ^* to be unity.

Let us now consider some examples.

4 Example (Multipliers are zero) The Lagrange Multiplier Theorem does not guarantee that all the multipliers on the constraints will be nonzero. In fact the multipliers on the constraints may all be zero. For instance consider the constrained maximum of

$$f(x, y) := -(x^2 + y^2)$$

subject to the single constraint

$$g(x,y) := y = 0.$$

Observe that $g'(x,y) = (0,1) \neq 0$, so the gradient is linearly independent. The point $(0,0)$ is a constrained maximizer of f, but $f'(x,y) = (-2x,-2y)$ is equal to zero at $(0,0)$. Thus the only way to solve $f'(0,0) + \lambda^* g'(0,0)$ is to set $\lambda^* = 0$. □

5 Example (Dependent constraint gradients) If you are like me, you may be tempted to think that if the gradients of the constraints are linearly dependent, then one of them may be redundant. This is not true. Consider the constrained maximum of

$$f(x,y) := x$$

subject to the two constraints

$$\begin{aligned} g_1(x,y) &:= y - x^2 = 0 \\ g_2(x,y) &:= y + x^2 = 0. \end{aligned}$$

It is easy to see that $(0,0)$ is the only point satisfying both constraints, and

$$g_1'(0,0) = (0,1) = g_2'(0,0).$$

Thus the gradients of the constraints are dependent at the maximizer. Since $f' = (1,0)$, there is no solution to $f'(0,0) + \lambda_1^* g_1(0,0) + \lambda_2^* g_2(0,0)$. There is however a nonzero solution to $\lambda_0^* f'(0,0) + \lambda_1^* g_1(0,0) + \lambda_2^* g_2(0,0)$, namely $\lambda_0^* = 0$, $\lambda_1^* = 1$, and $\lambda_2^* = -1$.

Notice that neither constraint is redundant, since if one of them is dropped, there are no constrained maxima. □

2 Second Order Necessary Conditions for a Constrained Maximum

6 Theorem (Necessary Second Order Conditions for a Maximum) *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ and let $x^* \in \operatorname{int} U$. Let $f, g_1, \ldots, g_m : U \to \boldsymbol{R}$ be C^2, and suppose x^* is a local constrained maximizer of f subject to $g(x) = 0$. Define the Lagrangean $L(x,\lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$. Assume that ${g_1}'(x^*), \ldots, {g_m}'(x^*)$ are linearly independent, so the conclusion of the Lagrange Multiplier Theorem holds, that is, there are $\lambda_1^*, \ldots, \lambda_m^*$ satisfying the first order conditions*

$$L_x'(x^*,\lambda^*) = f'(x^*) + \sum_{i=1}^m \lambda_i^* {g_i}'(x^*) = 0.$$

Then

$$\sum_{i=1}^n \sum_{j=1}^n D_{ij} L(x^*,\lambda^*) v_i v_j \leqslant 0,$$

for all $v \neq 0$ satisfying ${g_i}'(x^*) \cdot v = 0$, $i = 1, \ldots, m$.

3 Constrained Minimization

Since minimizing f is the same as maximizing $-f$, we do not need any new results for minimization, but there a few things worth pointing out.

The Lagrangean for maximizing $-f$ subject to $g_i = 0$, $i = 1, \dots, m$ is

$$-f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

The second order condition for maximizing $-f$ is that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \left(-D_{ij} f(x^*) + \sum_{i=1}^{m} \lambda^* D_{ij} g(x^*) \right) v_i v_j \leqslant 0,$$

for all $v \neq 0$ satisfying ${g_i}'(x^*) \cdot v = 0$, $i = 1, \dots, m$. This can be rewritten as

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \left(D_{ij} f(x^*) - \sum_{i=1}^{m} \lambda^* D_{ij} g(x^*) \right) v_i v_j \geqslant 0,$$

which suggests that it is more convenient to define the Lagrangean for a minimization problem as

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

The first order conditions will be exactly the same. For the second order conditions we have the following.

7 Theorem (Necessary Second Order Conditions for a Minimum) *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ and let $x^* \in \operatorname{int} U$. Let $f, g_1, \dots, g_m \colon U \to \boldsymbol{R}$ be C^2, and suppose x^* is a local constrained minimizer of f subject to $g(x) = 0$. Define the Lagrangean*

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

Assume that ${g_1}'(x^*), \dots, {g_m}'(x^*)$ are linearly independent, so the conclusion of the Lagrange Multiplier Theorem holds, that is, there are $\lambda_1^*, \dots, \lambda_m^*$ satisfying the first order conditions

$$L'_x(x^*, \lambda^*) = f'(x^*) - \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) = 0.$$

Then

$$\sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij} L(x^*, \lambda^*) v_i v_j \geqslant 0,$$

for all $v \neq 0$ satisfying ${g_i}'(x^*) \cdot v = 0$, $i = 1, \dots, m$.

4 Classical Envelope Theorem

8 Theorem (Envelope Theorem for Maximization) *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$ and $P \subset \boldsymbol{R}^{\ell}$ be open, and let $f \colon X \times P \to \boldsymbol{R}$ be C^1. For each $p \in P$, let $x^*(p)$ be an interior local maximizer of $f(\cdot, p)$. Assume that $x^* \colon P \to X$ is C^1. Set*

$$V(p) = f\left(x^*(p), p\right).$$

Then V is C^1 and

$$D_i V(p) = D_{n+i} f\left(x^*(p), p\right).$$

Or $\frac{\partial V}{\partial p_i} = \frac{\partial f}{\partial p_i} \big|_{x = x^*(p)}$.

Proof: By definition $V(p) \geqslant f(x,p)$ for all x, p, and $V(p) = f\left(x^*(p), p\right)$. Thus the function

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

achieves a minimum whenever $(p,x) = \big(p, x^*(p)\big)$. The first order conditions for a minimum of h are:

$$D_i h(p,x) = 0, \quad i = 1, \dots, n+\ell.$$

But $D_i h(p,x) = D_i V(p) - D_{n+i} f(x,p)$ for $i = 1, \dots, n$. ∎

9 Theorem (Envelope Theorem for Constrained Maximization) *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$ and $P \subset \boldsymbol{R}^{\ell}$ be open, and let $f, g_1, \dots, g_m \colon X \times P \to \boldsymbol{R}$ be C^1. For each $p \in P$, let $x^*(p)$ be an interior constrained local maximizer of $f(x,p)$ subject to $g(x,p) = 0$. Define the Lagrangean*

$$L(x, \lambda; p) = f(x,p) + \sum_{i=1}^{m} \lambda_i g_i(x,p),$$

and assume that the conclusion of the Lagrange Multiplier Theorem holds for each p, that is, there exist real numbers $\lambda_1^*(p), \dots, \lambda_m^*(p)$, such that the first order conditions Notation!!!!

$$\frac{\partial L\big(x^*(p), \lambda^*(p), p\big)}{\partial x} = f_x'\big(x^*(p), p\big) + \sum_{i=1}^{m} \lambda_i^*(p) {g_i}_x'\big(x^*(p), p\big) = 0$$

are satisfied. Assume that $x^* \colon P \to X$ and $\lambda^* \colon P \to \boldsymbol{R}^{\mathrm{m}}$ are C^1. Set

$$V(p) = f\left(x^*(p), p\right).$$

Then V is C^1 and

$$\frac{\partial V(p)}{\partial p_j} = \frac{\partial L\big(x^*(p), \lambda^*(p), p\big)}{\partial p_j} = \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^{m} \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j}.$$

Proof: Clearly V is C^1 as the composition of C^1 functions. Since x^* satisfies the constraints, we have

$$V(p) = f\left(x^*(p), p\right) = f\left(x^*(p), p\right) + \sum_{i=1}^{m} \lambda_i^*(p) g_i(x^*, p).$$

Therefore by the chain rule,

$$\begin{aligned}
\frac{\partial V(p)}{\partial p_j} &= \left(\sum_{k=1}^{n} \frac{\partial f(x^*, p)}{\partial x_k} \frac{\partial x^{*k}}{\partial p_j} \right) + \frac{\partial f(x^*, p)}{\partial p_j} \\
&\quad + \sum_{i=1}^{m} \left\{ \frac{\partial \lambda_i^*(p)}{\partial p_j} g_i(x^*, p) + \lambda^*(p) \left[\left(\sum_{k=1}^{n} \frac{\partial g_i(x^*, p)}{\partial x_k} \frac{\partial x^{*k}}{\partial p_j} \right) + \frac{\partial g_i(x^*, p)}{\partial p_j} \right] \right\} \\
&= \frac{\partial f(x^*, p)}{\partial p_j} + \sum_{i=1}^{m} \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j} \\
&\quad + \sum_{i=1}^{m} \frac{\partial \lambda_i^*(p)}{\partial p_j} g_i(x^*, p) \qquad (1) \\
&\quad + \sum_{k=1}^{n} \left(\frac{\partial f(x^*, p)}{\partial x_k} + \sum_{i=1}^{m} \lambda^*(p) \frac{\partial g_i(x^*, p)}{\partial x_k} \right) \frac{\partial x^{*k}}{\partial p_j}. \qquad (2)
\end{aligned}$$

The theorem now follows from the fact that both terms (1) and (2) are zero. Term (1) is zero since x^* satisfies the constraints, and term (2) is zero, since the first order conditions imply that each $\frac{\partial f(x^*,p)}{\partial x_k} + \sum_{i=1}^{m} \lambda^*(p) \frac{\partial g_i(x^*,p)}{\partial x_k} = 0$. ∎

10 Theorem (Envelope Theorem for Minimization) *Let $X \subset \boldsymbol{R}^{\mathrm{n}}$ and $P \subset \boldsymbol{R}^{\ell}$ be open, and let $f, g_1, \dots, g_m \colon X \times P \to \boldsymbol{R}$ be C^1. For each $p \in P$, let $x^*(p)$ be an interior constrained local maximizer of $f(x,p)$ subject to $g(x,p) = 0$. Define the Lagrangean*

$$L(x, \lambda; p) = f(x,p) - \sum_{i=1}^{m} \lambda_i g_i(x,p),$$

and assume that the conclusion of the Lagrange Multiplier Theorem holds for each p, that is, there exist real numbers $\lambda_1^*(p), \dots, \lambda_m^*(p)$, such that the first order conditions Notation!!!!

$$\frac{\partial L\big(x^*(p), \lambda^*(p), p\big)}{\partial x} = f'_x\big(x^*(p), p\big) - \sum_{i=1}^{m} \lambda_i^*(p) {g_i}'_x\big(x^*(p), p\big) = 0$$

are satisfied. Assume that $x^* \colon P \to X$ and $\lambda^* \colon P \to \boldsymbol{R}^{\mathrm{m}}$ are C^1. Set

$$V(p) = f\left(x^*(p), p\right).$$

Then V is C^1 and

$$\frac{\partial V(p)}{\partial p_j} = \frac{\partial L\big(x^*(p), \lambda^*(p), p\big)}{\partial p_j} = \frac{\partial f(x^*, p)}{\partial p_j} - \sum_{i=1}^{m} \lambda_i^*(p) \frac{\partial g_i(x^*, p)}{\partial p_j}.$$

The proof is the same as that of Theorem 9.

5 Inequality constraints

The classical Lagrange Multiplier Theorem deals only with equality constraints. Now we take up inequality constraints. I learned this approach from Quirk [9].

11 Theorem *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ be open, and let $f, g_1, \dots, g_m \colon U \to \boldsymbol{R}$ be twice continuously differentiable on U. Let x^* be a constrained local maximizer of f subject to $g(x) \geqq 0$ and $x \geqq 0$.*

Let $B = \{i : g_i(x^*) = 0\}$, the set of binding constraints, and let $Z = \{j : x_j = 0\}$, the set of binding nonnegativity constraints. Assume that $\{{g_i}'(x^*) : i \in B\} \cup \{e^j : j \in Z\}$ is linearly independent. Then there exists $\lambda^* \in \boldsymbol{R}^{\mathrm{m}}$ such that

$$f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \leqq 0. \tag{3}$$

$$x^* \cdot \left(f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \right) = 0 \tag{4}$$

$$\lambda^* \geqq 0. \tag{5}$$

$$\lambda^* \cdot g(x^*) = 0. \tag{6}$$

Proof: Introduce $m + n$ **slack variables** $y_1, \dots, y_m$ and $z_1, \dots, z_n$, and consider the equality constrained maximization problem:

maximize $\bar{f}(x)$ subject to $g^i(x) - y_i^2 = 0$, $i = 1, \dots, m$, and $x_j - z_j^2 = 0$, $j = 1, \dots, n$.

Define y^* and z^* by

$$y_i^* = \sqrt{g_i(x^*)} \tag{7}$$

$$z_j^* = \sqrt{x_j^*}, \tag{8}$$

Observe that (x^*, y^*, z^*) solves the equality constrained maximization problem.

So on $U \times \boldsymbol{R}^{\mathrm{m}} \times \boldsymbol{R}^{\mathrm{n}}$ define

$$\begin{aligned}\bar{f}(x,y,z) &= f(x),\\ \bar{g}_i(x,y,z) &= g_i(x) - y_i^2, \quad i = 1,\dots,m,\\ \bar{h}_j(x,y,z) &= x_j - z_j^2, \quad j = 1,\dots,n.\end{aligned}$$

Note that these functions are also twice continuously differentiable. Then (x^*, y^*, z^*) solves the revised equality constrained maximization problem:

maximize $\bar{f}(x,y,z)$ subject to $\bar{g}_i(x,y,z) = 0$, $i = 1,\dots,m$, and $\bar{h}_j(x,y,z) = 0$, $j = 1,\dots,n$.

In order to apply the Lagrange Multiplier Theorem to this revised equality constrained problem, we need to verify that the gradients of $\bar{g}_i'(x^*, y^*, z^*)$, $i = 1,\dots,m$ and $\bar{h}_j'(x^*, y^*, z^*)$, $j = 1,\dots,n$ of the constraints with respect to the variables x, y, z are linearly independent. So suppose $\sum_{i=1}^m \alpha_i \bar{g}_i' + \sum_{j=1} n\beta_j \bar{h}_j' = 0$. Now

$$\bar{g}_i'(x^*, y^*, z^*) = \begin{bmatrix} g_i{}'(x^*) \\ -2y_i^* e^i \\ 0 \end{bmatrix} \qquad \text{and} \qquad \bar{h}_j'(x^*, y^*, z^*) = \begin{bmatrix} e^j \\ 0 \\ -2z_j^* e^j \end{bmatrix}. \tag{9}$$

So the y_i component of this sum is just $-2\alpha_i y_i^*$. Therefore

$$i \notin B \iff y_i^* > 0 \implies \alpha_i = 0.$$

Similarly the z_j component is $-2\beta_j z_j^*$, so

$$j \notin Z \iff z_j^* > 0 \implies \beta_j = 0.$$

Given this, the x component is just

$$\sum_{i \in B} \alpha_i \bar{g}_i'(x^*) + \sum_{j \in Z} \beta_j e^j = 0.$$

By hypothesis, these vectors are linearly independent, so we conclude that $\alpha_i = 0$, $i \in B$, and $\beta_j = 0$, $j \in Z$, which proves the linear independence of the constraint gradients of the revised equality problem.

So form the Lagrangean

$$\begin{aligned}\bar{L}(x,y,z;\lambda,\mu) &= \bar{f}(x,y,z) + \sum_{i=1}^m \lambda_i \bar{g}_i(x,y,x) + \sum_{j=1}^n \mu_j \bar{h}_j(x,y,z).\\ &= f(x) + \sum_{i=1}^m \lambda_i \big(g_i(x) - y_i^2\big) + \sum_{j=1}^n \mu_j (x_j - z_j^2).\end{aligned}$$

Then by the Lagrange Multiplier Theorem 2 there are multipliers λ_i^*, $i = 1,\dots,m$ and μ_j^*, $j = 1,\dots,n$ such that the following first order conditions are satisfied.

$$\frac{\partial f(x^*)}{\partial x_j} + \sum_{i=1}^m \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} + \mu_j^* = 0 \qquad j = 1,\dots,n, \tag{10}$$

$$-2\lambda_1^* y_i^* = 0 \qquad i = 1,\dots,m, \tag{11}$$

$$-2\mu_j^* z_j^* = 0 \qquad j = 1,\dots,n. \tag{12}$$

Now the Hessian of the Lagrangean $\bar{L}$ (with respect to (x, y, z) and evaluated at the point $(x^*, y^*, z^*; \lambda^*, \mu^*)$) is block diagonal:

$$\left[\begin{array}{ccc|ccc|ccc}
\frac{\partial^2 L}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} & & & & & & \\
\vdots & & \vdots & & & & & & \\
\frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_n} & & & & & & \\
\hline
 & & & -2\lambda_1^* & & & & & \\
 & & & & \ddots & & & & \\
 & & & & & -2\lambda_m^* & & & \\
\hline
 & & & & & & -2\mu_1^* & & \\
 & & & & & & & \ddots & \\
 & & & & & & & & -2\mu_n^*
\end{array}\right]$$

From the second order conditions (Theorem 6) for the revised equality constrained problem, we know that this Hessian is negative semidefinite under constraint. That is, if a vector is orthogonal to the gradients of the constraints, then the quadratic form in the Hessian is nonpositive. In particular, consider a vector of the form $v = (0, e^k, 0) \in \boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}} \times \boldsymbol{R}^{\mathrm{n}}$. It follows from (9) that for $i \neq k$ this vector v is orthogonal to $\bar{g}_i'(x^*, y^*, z^*)$, and also orthogonal to $\bar{h}_j'(x^*, y^*, z^*)$, $j = 1, \ldots, n$. The vector v is orthogonal to $\bar{g}_k'(x^*, y^*, z^*)$ if and only if $y_k^* = 0$, that is, when $k \in B$. Thus for $k \in B$ the second order conditions imply $-2\lambda_k^* \leqslant 0$, so

$$g_i(x^*) = 0 \implies \lambda_i^* \geqslant 0.$$

Next consider a vector of the form $u = (0, 0, e^k) \in \boldsymbol{R}^{\mathrm{n}} \times \boldsymbol{R}^{\mathrm{m}} \times \boldsymbol{R}^{\mathrm{n}}$. It follows from (9) that this vector u is orthogonal to each $\bar{g}_i'(x^*, y^*, z^*)$ each $\bar{h}_j'(x^*, y^*, z^*)$ for $j \neq k$. The vector u is orthogonal to $\bar{h}_k'(x^*, y^*, z^*)$ if and only if $z_k^* = 0$, that is, $j \in Z$. Again the second order conditions imply that the quadratic form in u, which has value $-2\mu_k^*$ is nonnegative for $k \in Z$, so

$$x_j^* = 0 \implies \mu_j^* \geqslant 0.$$

Now if $i \notin B$, that is, $g_i(x^*) > 0$, so that $y_i^* > 0$, then from the first order condition (11) we have $\lambda_i^* = 0$. Also, from (12), if $x_j^* > 0$, so that $z_j^* > 0$, then $\mu_j^* = 0$. That is,

$$g_i(x^*) > 0 \implies \lambda_i^* = 0 \qquad \text{and} \qquad x_j^* > 0 \implies \mu_j^* = 0$$

Combining this with the paragraph above we see that $\lambda^* \geqq 0$ and $\mu^* \geqq 0$. Thus (10) implies conclusion (3). A little more thought will show you that we have just deduced conditions (4) through (6) as well. ∎

There is a simple variation on the slack variable approach that applies to mixed inequality and equality constraints. To prove the next result, simply omit the slack variables for the equality constraints and follow the same proof as in Theorem 11.

12 Corollary *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ be open, and let $f, g_1, \ldots, g_m \colon U \to \boldsymbol{R}$ be twice continuously differentiable on U. Let x^* be a constrained local maximizer of f subject to*

$$\begin{aligned} g_i(x) &= 0 \quad i \in E, \\ g_i(x) &\geqslant 0 \quad i \in E^c, \\ x_j &\geqslant 0 \quad j \in N. \end{aligned}$$

Let $B = \{i \in E^c : g_i(x^*) = 0\}$ (binding inequality constraints), and let $Z = \{j \in N : x_j = 0\}$ (binding nonnegativity constraints). Assume that

$$\{{g_i}'(x^*) : i \in E \cup B\} \cup \{e^j : j \in Z\} \text{ is linearly independent,}$$

then there exists $\lambda^* \in \boldsymbol{R}^{\mathrm{m}}$ such that

$$\begin{aligned} \frac{\partial f(x^*)}{\partial x_j} + \textstyle\sum_{j=1}^m \lambda_i^* \frac{\partial g_j(x^*)}{\partial x_j} &\leqslant 0 \quad j \in N, \\ \frac{\partial f(x^*)}{\partial x_i} + \textstyle\sum_{j=1}^m \lambda_i^* \frac{\partial g_j(x^*)}{\partial x_i} &= 0 \quad j \in N^c, \\ \lambda_i^* &\geqslant 0 \quad i \in E^c. \\ x^* \cdot \Big(f'(x^*) + \textstyle\sum_{j=1}^m \lambda_i^* {g_j}'(x^*) \Big) &= 0 \\ \lambda^* \cdot g(x^*) &= 0. \end{aligned}$$

We now translate the result for minimization.

13 Theorem (Minimization) *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ be open, and let $f, g_1, \ldots, g_m : U \to \boldsymbol{R}$ be twice continuously differentiable on U. Let x^* be a constrained local minimizer of f subject to $g(x) \geqq 0$ and $x \geqq 0$.*

Let $B = \{i : g_i(x^*) = 0\}$, the set of binding constraints, and let $Z = \{j : x_j = 0\}$, the set of binding nonnegativity constraints. Assume that $\{{g_i}'(x^*) : i \in B\} \cup \{e^j : j \in Z\}$ is linearly independent. Then there exists $\lambda^* \in \boldsymbol{R}^{\mathrm{m}}$ such that

$$f'(x^*) - \sum_{i=1}^m \lambda_i^* {g_i}'(x^*) \geqq 0. \tag{13}$$

$$x^* \cdot \left(f'(x^*) - \sum_{i=1}^m \lambda_i^* {g_i}'(x^*) \right) = 0 \tag{14}$$

$$\lambda^* \geqq 0. \tag{15}$$

$$\lambda^* \cdot g(x^*) = 0. \tag{16}$$

Proof: As in the proof of Theorem 11, introduce $m+n$ slack variables $y_1, \ldots, y_m$ and $z_1, \ldots, z_n$, and define $\bar{f}(x, y, z) = f(x)$, $\bar{g}_i(x, y, z) = g_i(x) - y_i^2$, $i = 1, \ldots, m$ and $\bar{h}_j(x, y, z) = x_j - z_j^2$, $j = 1, \ldots, n$. Again define y^* by $y_i^* = \sqrt{g_i(x^*)}$ and z^* by $z_j^* = \sqrt{x_j^*}$. Observe that (x^*, y^*, z^*) solves the revised equality constrained minimization problem:

minimize $\bar{f}(x, y, z)$ subject to $\bar{g}_i(x, y, z) = 0$, $i = 1, \ldots, m$, and $\bar{h}_j(x, y, z) = 0$, $j = 1, \ldots, n$.

The proof of the linear independence of the constraint gradients of the revised equality problem is the same as in Theorem 11.

So form the Lagrangean

$$\bar{L}(x, y, z; \lambda, \mu) = f(x) - \sum_{i=1}^m \lambda_i \big(g_i(x) - y_i^2 \big) - \sum_{j=1}^n \mu_j (x_j - z_j^2).$$

Then by the Lagrange Multiplier Theorem 2 there are multipliers λ_i^*, $i = 1, \ldots, m$ and μ_j^*, $j = 1, \ldots, n$ such that the following first order conditions are satisfied.

$$\frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} - \mu_j^* = 0 \qquad j = 1, \ldots, n, \tag{17}$$

$$2\lambda_1^* y_i^* = 0 \qquad i = 1, \ldots, m, \tag{18}$$

$$2\mu_j^* z_j^* = 0 \qquad j = 1, \ldots, n. \tag{19}$$

The Hessian of the Lagrangean $\bar{L}$ (with respect to (x, y, z) and evaluated at $(x^*, y^*, z^*; \lambda^*, \mu^*)$) is:

$$\left[\begin{array}{ccc|ccc|ccc}
\frac{\partial^2 L}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} & & & & & & \\
\vdots & & \vdots & & & & & & \\
\frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_n} & & & & & & \\
\hline
 & & & 2\lambda_1^* & & & & & \\
 & & & & \ddots & & & & \\
 & & & & & 2\lambda_m^* & & & \\
\hline
 & & & & & & 2\mu_1^* & & \\
 & & & & & & & \ddots & \\
 & & & & & & & & 2\mu_n^*
\end{array}\right]$$

From the second order conditions for minimization (Theorem 7) for the revised equality constrained problem, we know that this Hessian is positive semidefinite under constraint. In particular, as in the proof of Theorem 11, we have that

$$g_i(x^*) = 0 \implies \lambda_i^* \geqslant 0.$$

$$x_j^* = 0 \implies \mu_j^* \geqslant 0.$$

From the first order conditions, if $i \notin B$, that is, $g_i(x^*) > 0$, so that $y_i^* = 0$, then $\lambda_i^* = 0$. Also if $x_j^* > 0$, so that $z_j^* = 0$, then $\mu_j^* = 0$. That is,

$$g_i(x^*) > 0 \implies \lambda_i^* = 0 \qquad \text{and} \qquad x_j^* > 0 \implies \mu_j^* = 0$$

Combining this with the paragraph above we see that $\lambda^* \geqq 0$ and $\mu^* \geqq 0$. Thus (17) implies conclusion (13). A little more thought will show you that we have just deduced conditions (14) through (16) as well. ∎

14 Corollary *Let $U \subset \boldsymbol{R}^{\mathrm{n}}$ be open, and let $f, g_1, \ldots, g_m \colon U \to \boldsymbol{R}$ be twice continuously differentiable on U. Let x^* be a constrained local minimizer of f subject to*

$$g_i(x) = 0 \quad i \in E,$$

$$g_i(x) \geqslant 0 \quad i \in E^c,$$

$$x_j \geqslant 0 \quad j \in N.$$

Let $B = \{i \in E^c : g_i(x^*) = 0\}$ (binding inequality constraints), and let $Z = \{j \in N : x_j = 0\}$ (binding nonnegativity constraints). Assume that

$$\{g_i{}'(x^*) : i \in E \cup B\} \cup \{e^j : j \in Z\} \text{ is linearly independent,}$$

then there exists $\lambda^* \in \boldsymbol{R}^{\mathrm{m}}$ such that

$$\begin{aligned}
\frac{\partial f(x^*)}{\partial x_j} - \textstyle\sum_{j=1}^m \lambda_i^* \frac{\partial g_j(x^*)}{\partial x_j} &\geqslant 0 \quad j \in N, \\
\frac{\partial f(x^*)}{\partial x_i} + \textstyle\sum_{j=1}^m \lambda_i^* \frac{\partial g_j(x^*)}{\partial x_i} &= 0 \quad j \in N^c, \\
\lambda_i^* &\geqslant 0 \quad i \in E^c. \\
x^* \cdot \Big(f'(x^*) + \textstyle\sum_{j=1}^m \lambda_i^* {g_j}'(x^*) \Big) &= 0 \\
\lambda^* \cdot g(x^*) &= 0.
\end{aligned}$$

6 Karush–Kuhn–Tucker Theory

A drawback of the slack variable approach is that it assumes twice continuous differentiability in order to apply the second order conditions and thus conclude $\lambda^* \geqq 0$ and $\mu^* \geqq 0$. Fortunately, Karush [5] and Kuhn and Tucker [7] provide another approach that remedies this defect. They only assume differentiability, and replace the independence condition on gradients by a weaker but more obscure condition called the Karush–Kuhn–Tucker Constraint Qualification.

15 Definition *Let $f, g_1, \ldots, g_m \colon \boldsymbol{R}^{\mathrm{n}}_+ \to \boldsymbol{R}$. Let*

$$C = \{x \in \boldsymbol{R}^{\mathrm{n}} : x \geqq 0,\ g_i(x) \geqslant 0,\ i = 1, \ldots, m\}.$$

In other words, C is the constraint set. Consider a point $x^* \in C$ and define

$$B = \{i : g_i(x^*) = 0\} \text{ and } Z = \{j : x_j = 0\},$$

the set of binding constraints and binding nonnegativity constraints, respectively. The point x^* satisfies the ***Karush–Kuhn–Tucker Constraint Qualification*** *if $f, g_1, \ldots, g_m$ are differentiable at x^*, and for every $v \in \boldsymbol{R}^{\mathrm{n}}$ satisfying*

$$\begin{aligned}
v_j = v \cdot e^j &\geqslant 0 \qquad j \in Z, \\
v \cdot {g_i}'(x^*) &\geqslant 0 \qquad i \in B,
\end{aligned}$$

there is a continuous curve $\xi \colon [0, \varepsilon) \to \boldsymbol{R}^{\mathrm{n}}$ satisfying

$$\begin{aligned}
\xi(0) &= x^*, \\
\xi(t) \in C & \quad \text{for all } t \in [0, \varepsilon), \\
D\xi(0) &= v,
\end{aligned}$$

where $D\xi(0)$ is the one-sided directional derivative at 0. Consistent notation?

This condition is actually a little weaker than Kuhn and Tucker's condition. They assumed that the functions $f, g_1, \ldots, g_m$ were differentiable everywhere and required ξ to be differentiable everywhere. You can see that it may be difficult to verify it in practice.

16 Theorem (Karush–Kuhn–Tucker) *Let $f, g_1, \ldots, g_m \colon \boldsymbol{R}^{\mathrm{n}}_+ \to \boldsymbol{R}$ be differentiable at x^*, and let x^* be a constrained local maximizer of f subject to $g(x) \geqq 0$ and $x \geqq 0$.*

Let $B = \{i : g_i(x^*) = 0\}$, the set of binding constraints, and let $Z = \{j : x_j^* = 0\}$, the set of binding nonnegativity constraints. Assume that x^* satisfies the Karush–Kuhn–Tucker Constraint Qualification. Then there exists $\lambda^* \in \boldsymbol{R}^{\mathrm{m}}$ such that

$$f'(x^*) + \sum_{i=1}^m \lambda_i^* {g_i}'(x^*) \leqq 0,$$

$$x^* \cdot \left(f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \right) = 0,$$

$$\lambda^* \geqq 0,$$

$$\lambda^* \cdot g(x^*) = 0.$$

To better understand the hypotheses of the theorem, let's look at a classic example of its failure (cf. Kuhn and Tucker [7]).

17 Example (Failure of the Karush–Kuhn–Tucker Constraint Qualification) Consider the functions $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ via $f(x, y) = x$ and $g : \boldsymbol{R}^2 \to \boldsymbol{R}$ via $g(x, y) = (1 - x)^3 - y$. The curve $g = 0$ is shown in Figure 3, and the constraint set in Figure 4.

Clearly $(x^*, y^*) = (1, 0)$ maximizes f subject to $(x, y) \geqq 0$ and $g \geqslant 0$. At this point we have $g'(1, 0) = (0, -1)$ and $f' = (1, 0)$ everywhere. Note that no λ (nonnegative or not) satisfies

$$(1, 0) + \lambda(0, -1) \leqq (0, 0).$$

Fortunately for the theorem, the Constraint Qualification fails at $(1, 0)$. To see this, note that the constraint $g \geqslant 0$ binds, that is $g(1, 0) = 0$ and the second coordinate of (x^*, y^*) is zero. Suppose $v = (v_x, v_y)$ satisfies

$$v \cdot g'(1, 0) = v \cdot (0, -1) = -v_y \leqslant 0 \qquad \text{and} \qquad v \cdot e^2 = v_y \geqslant 0,$$

that is, $v_y = 0$. For instance, take $v = (1, 0)$. The constraint qualification requires that there is a path starting at $(1, 0)$ in the direction $(1, 0)$ that stays in the constraint set. Clearly no such path exists, so the constraint qualification fails. □

7 Karush–Kuhn–Tucker Theorem for Minimization

18 Theorem (Karush–Kuhn–Tucker) *Let $f, g_1, \ldots, g_m : \boldsymbol{R}^n_+ \to \boldsymbol{R}$ be differentiable at x^*, and let x^* be a constrained local minimizer of f subject to $g(x) \geqq 0$ and $x \geqq 0$.*

Let $B = \{i : g_i(x^*) = 0\}$, the set of binding constraints, and let $Z = \{j : x_j = 0\}$, the set of binding nonnegativity constraints. Assume that x^* satisfies the Karush–Kuhn–Tucker Constraint Qualification. Then there exists $\lambda^* \in \boldsymbol{R}^m$ such that

$$f'(x^*) - \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \geqq 0,$$

$$x^* \cdot \left(f'(x^*) - \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \right) = 0,$$

$$\lambda^* \geqq 0,$$

$$\lambda^* \cdot g(x^*) = 0.$$

Proof: Minimizing f is the same as maximizing $-f$. The Karush–Kuhn–Tucker conditions for this imply that there exists $\lambda^* \in \boldsymbol{R}^m_+$ such that

$$-f'(x^*) + \sum_{i=1}^{m} \lambda_i^* {g_i}'(x^*) \leqq 0,$$

and the conclusion follows by multiplying this by -1. ■

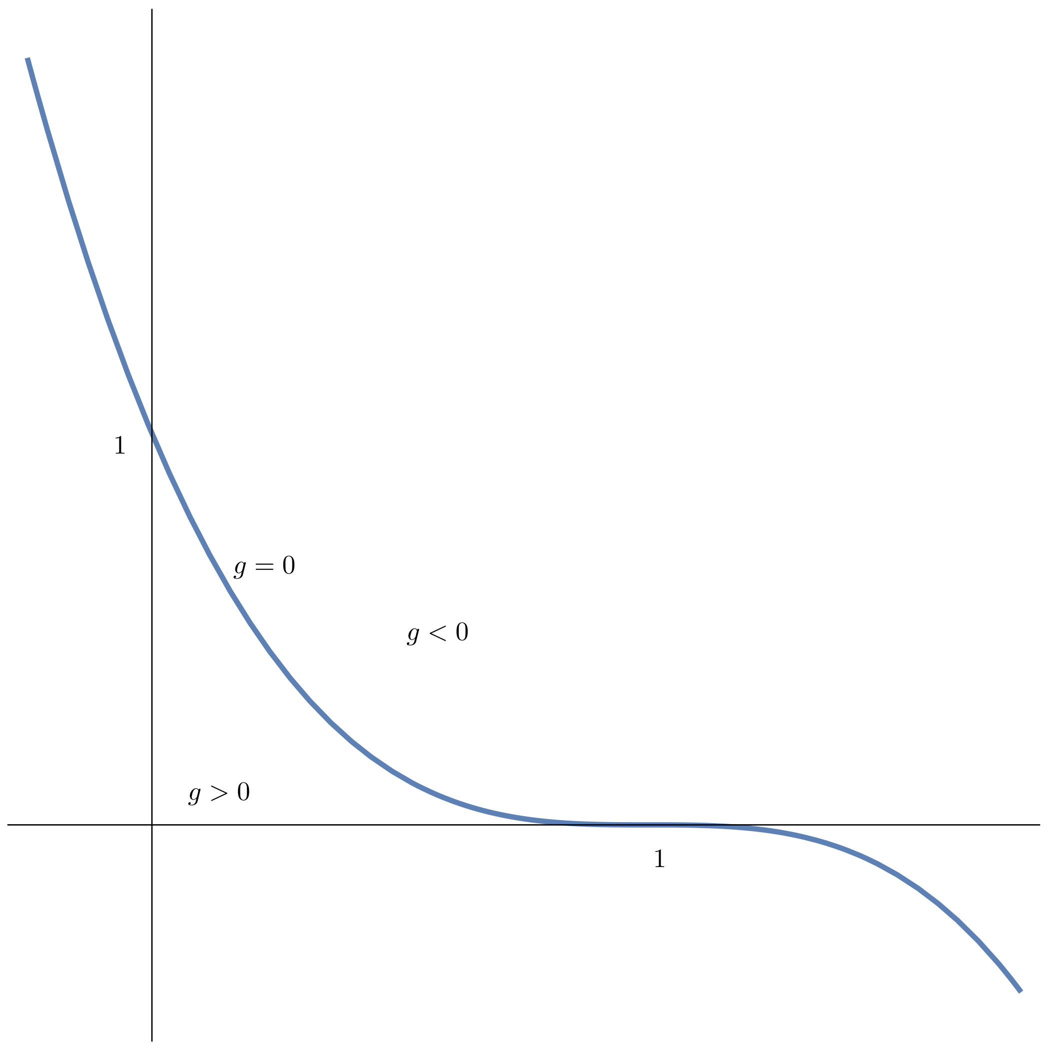

Figure 3. The function $g(x, y) = (1 - x)^3 - y$.

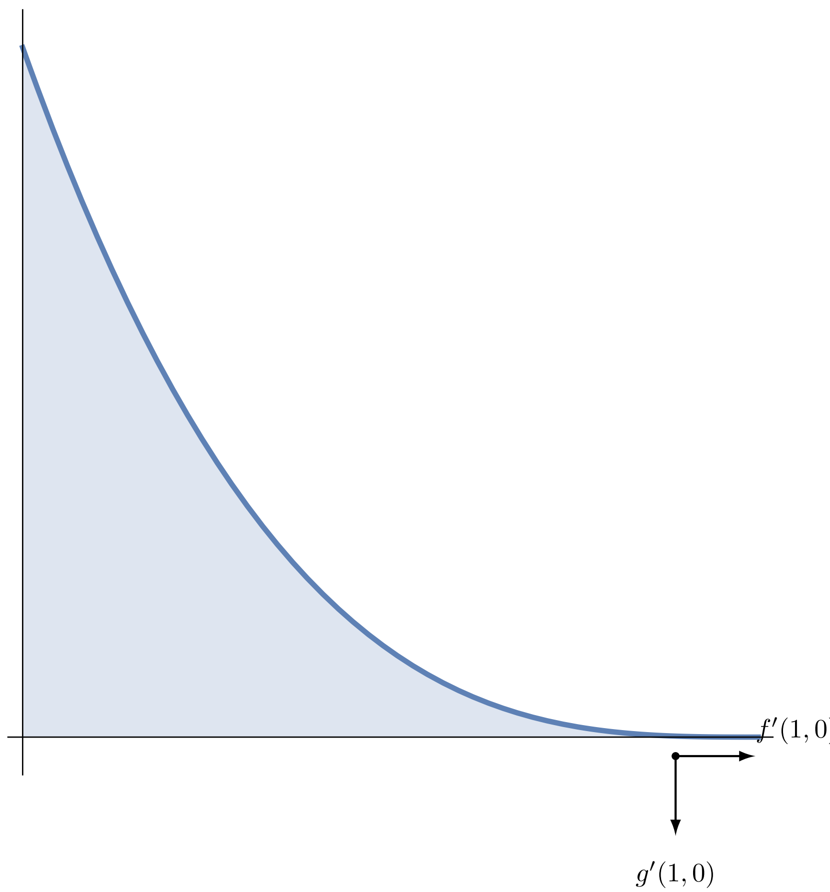

Figure 4. This constraint set violates the Constraint Qualification. (Note: f' and g' are not to scale.)

8 Quasiconcave functions

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

19 Definition *A function $f\colon C \to \boldsymbol{R}$ on a convex subset C of a vector space is:*

- ***quasiconcave*** *if*
$$f\big(\lambda x + (1-\lambda)y\big) \geqslant \min\{f(x), f(y)\}$$
for all x, y in C and all $0 \le \lambda \le 1$.

- ***explicitly quasiconcave*** *or* ***semistrictly quasiconcave*** *if it is quasiconcave and $f(x) > f(y)$ implies $f\big(\lambda x + (1-\lambda)y\big) > f(y)$ for every $\lambda \in (0,1)$.*

- ***strictly quasiconcave*** *if $f\big(\lambda x + (1-\lambda)y\big) > \min\{f(x), f(y)\}$ for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$.*

- ***quasiconvex*** *if $f\big(\lambda x + (1-\lambda)y\big) \le \max\{f(x), f(y)\}$ for all x, y in C and all $0 \le \lambda \le 1$.*

- ***explicitly quasiconvex*** *or* ***semistrictly quasiconvex*** *if it is quasiconvex and $f(x) < f(y)$ implies $f\big(\lambda x + (1-\lambda)y\big) < f(y)$ for every $\lambda \in (0,1)$.*

- ***strictly quasiconvex*** *if $f\big(\lambda x + (1-\lambda)y\big) < \max\{f(x), f(y)\}$ for all x, y in C with $x \neq y$ and all $0 < \lambda < 1$.*

Then next lemma is a simple consequence of the definitions.

20 Lemma *A concave function is quasiconcave. A convex function is quasiconvex.*

Of course, not every quasiconcave function is concave. For instance, the function in Figure 5 is quasiconcave but not concave.

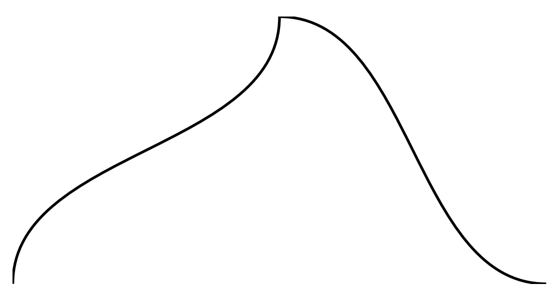

Figure 5. This function is quasiconcave, but not concave.

Characterizations of quasiconcavity are given in the next lemma.

21 Lemma *For a function $f\colon C \to \boldsymbol{R}$ on a convex set, the following are equivalent:*

1. *The function f is quasiconcave.*
2. *For each $\alpha \in \boldsymbol{R}$, the strict upper contour set $[f(x) > \alpha]$ is convex, but possibly empty.*
3. *For each $\alpha \in \boldsymbol{R}$, 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 $0 \leqslant \lambda \leqslant 1$ we have

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

(2) $\implies$ (3) Note that

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

and recall that the intersection of convex sets is convex.

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

Quasiconcavity is a weaker property than concavity. For instance, the sum of two quasiconcave functions need not be quasiconcave.

22 Example (Sum of quasiconcave functions is not quasiconcave) Define f and g on the real line by

$$f(x) = \begin{cases} 0 & x \leqslant 0 \\ x & x \geqslant 0, \end{cases} \qquad \text{and} \qquad g(x) = \begin{cases} -x & x \leqslant 0 \\ 0 & x \geqslant 0. \end{cases}$$

Then both f and g are quasiconcave. For instance, $[f \geqslant \alpha]$ is either $\boldsymbol{R}$ or $[\alpha, \infty)$, each of which is a convex set. (Observe that f and g are both convex functions as well!) The sum $(f+g)(x) = |x|$ is not quasiconcave, since for $\alpha > 0$, $[f + g \geqslant \alpha] = (-\infty, -\alpha] \cup [\alpha, \infty)$, which is not a convex set. □

The next result has applications to production functions. (Cf. Jehle [4, Theorem 5.2.1, pp. 224–225] and Shephard [10, pp. 5–7].)

23 Theorem *Let* $f\colon \boldsymbol{R}^{\mathrm{n}}_{+} \to \boldsymbol{R}_{+}$ *be nonnegative, nondecreasing, quasiconcave, and positively homogeneous of degree* k *where* $0 < k \leqslant 1$. *Then* f *is concave.*

Proof: Let $x, y \in \boldsymbol{R}^{\mathrm{n}}$ and suppose first that $f(x) = \alpha > 0$ and $f(y) = \beta > 0$. Then by homogeneity,

$$f\left(\frac{x}{\alpha^{\frac{1}{k}}}\right) = f\left(\frac{y}{\beta^{\frac{1}{k}}}\right) = 1$$

By quasiconcavity,

$$f\left(\lambda \frac{x}{\alpha^{\frac{1}{k}}} + (1-\lambda)\frac{y}{\beta^{\frac{1}{k}}}\right) \geqslant 1$$

for $0 \leqslant \lambda \leqslant 1$. So setting $\lambda = \frac{\alpha^{\frac{1}{k}}}{\alpha^{\frac{1}{k}} + \beta^{\frac{1}{k}}}$, we have

$$f\left(\frac{x}{\alpha^{\frac{1}{k}} + \beta^{\frac{1}{k}}} + \frac{y}{\alpha^{\frac{1}{k}} + \beta^{\frac{1}{k}}}\right) \geqslant 1.$$

By homogeneity,

$$f(x+y) \geqslant (\alpha^{\frac{1}{k}} + \beta^{\frac{1}{k}})^k = \big[f(x)^{\frac{1}{k}} + f(y)^{\frac{1}{k}}\big]^k. \tag{20}$$

Observe that since f is nonnegative and nondecreasing, (20) holds even if $f(x) = 0$ or $f(y) = 0$. Now replace x by μx and y by $(1-\mu)y$ in (20), where $0 \leqslant \mu \leqslant 1$, to get

$$\begin{aligned} f\big(\mu x + (1-\mu)y\big) &\geqslant \Big[f(\mu x)^{\frac{1}{k}} + f\big((1-\mu)y\big)^{\frac{1}{k}}\Big]^k \\ &= \Big[\mu f(x)^{\frac{1}{k}} + (1-\mu)f(y)^{\frac{1}{k}}\Big]^k \\ &\geqslant \mu\big(f(x)^{\frac{1}{k}}\big)^k + (1-\mu)\big(f(y)^{\frac{1}{k}}\big)^k \\ &= \mu f(x) + (1-\mu)f(y), \end{aligned}$$

where the last inequality follows from the concavity of $\gamma \mapsto \gamma^k$. Since x and y are arbitrary, f is concave. ∎

9 Quasiconcavity and Differentiability

Quasiconcavity has implications for derivatives.

24 Proposition *Let $C \subset \boldsymbol{R}^{\mathrm{n}}$ be convex and let $f\colon C \to \boldsymbol{R}$ be quasi-concave. Let y belong to C and assume that f has a one-sided directional derivative $f'(x; y-x) = \lim_{\lambda \downarrow 0} \frac{f\left(x+\lambda(y-x)\right)-f(x)}{\lambda}$. Then* Notation? Definition?

$$f(y) \geqslant f(x) \quad \Longrightarrow \quad f'(x; y-x) \geqslant 0.$$

In particular, if f is differentiable at x, then $f'(x)\cdot(y-x) \geqslant 0$ whenever $f(y) \geqslant f(x)$.

Proof: If $f(y) \geqslant f(x)$, then $f\big(x+\lambda(y-x)\big) = f\big((1-\lambda)x + \lambda y\big) \geqslant f(x)$ for $0 < \lambda \leqslant 1$ by quasiconcavity. Rearranging implies $\frac{f\left(x+\lambda(y-x)\right)-f(x)}{\lambda} \geqslant 0$ and taking limits gives the desired result. ∎

25 Theorem *Let $C \subset \boldsymbol{R}^{\mathrm{n}}$ be open and let $f\colon C \to \boldsymbol{R}$ be quasiconcave and twice-differentiable at $x \in C$. Then*

$$\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x)}{\partial x_i x_j} v_i v_j \leqslant 0 \qquad \text{for any } v \text{ satisfying } f'(x)\cdot v = 0.$$

Proof: Pick $v \in \boldsymbol{R}^{\mathrm{n}}$ and define $g(\lambda) = f(x+\lambda v)$. Then $g(0) = f(x)$, $g'(0) = f'(x)\cdot v$, and $g''(0) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(x)}{\partial x_i \partial x_j} v_i v_j$. What we have to show is that if $g'(0) = 0$, then $g''(0) \leqslant 0$. Assume for the sake of contradiction that $g'(0) = 0$ and $g''(0) > 0$. Then g has a strict local minimum at zero. That is, for $\varepsilon > 0$ small enough, $f(x+\varepsilon v) > f(x)$ and $f(x-\varepsilon v) > f(x)$. But by quasiconcavity,

$$f(x) = f\big(\tfrac{1}{2}(x+\varepsilon v) + \tfrac{1}{2}(x-\varepsilon v)\big) \geqslant \min\{f(x+\varepsilon v), f(x-\varepsilon v)\} > f(x),$$

a contradiction. ∎

10 Quasiconcavity and First Order Conditions

The following theorem and its proof may be found in Arrow and Enthoven [1].

26 Theorem (Arrow–Enthoven) *Let $f, g_1, \ldots, g_m \colon \boldsymbol{R}^n_+ \to \boldsymbol{R}$ be differentiable and quasi-concave. Suppose $x^* \in \boldsymbol{R}^n_+$ and $\lambda^* \in \boldsymbol{R}^m$ satisfy the constraints $g(x^*) \geqq 0$ and $x^* \geqq 0$ and the Karush–Kuhn–Tucker–Lagrange first order conditions:*

$$f'(x^*) + \textstyle\sum_{j=1}^m \lambda_i^* {g_j}'(x^*) \leqq 0$$

$$x^* \cdot \Big(f'(x^*) + \textstyle\sum_{j=1}^m \lambda_i^* {g_j}'(x^*)\Big) = 0$$

$$\lambda^* \geqq 0$$

$$\lambda^* \cdot g(x^*) = 0.$$

*Say that a variable x_j is **relevant** if it may take on a strictly positive value in the constraint set. That is, if there exists some $\hat{x} \geqq 0$ satisfying $\hat{x}_j > 0$ and $g(\hat{x}) \geqq 0$.*

Suppose one of the following conditions is satisfied:

1. $\dfrac{\partial f(x^*)}{\partial x_{j_0}} < 0$ *for some relevant variable x_{j_0}.*
2. $\dfrac{\partial f(x^*)}{\partial x_{j_1}} > 0$ *for some relevant variable x_{j_1}.*
3. $f'(x^*) \neq 0$ *and f is twice differentiable in a neighborhood of x^*.*
4. f *is concave.*

Then x^* maximizes $f(x)$ subject to the constraints $g(x) \geqq 0$ and $x \geqq 0$.

References

[1] K. J. Arrow and A. C. Enthoven. 1961. Quasi-concave programming. *Econometrica* 29(4):779–800. http://www.jstor.org/stable/1911819

[2] K. C. Border. Notes on the implicit function theorem. http://www.its.caltech.edu/~kcborder/Notes/IFT.pdf

[3] C. Carathéodory. 1982. *Calculus of variations*, 2d. ed. New York: Chelsea. This was originally published in 1935 in two volumes by B. G. Teubner in Berlin as *Variationsrechnung und Partielle Differentialgleichungen erster Ordnung*. In 1956 the first volume was edited and updated by E. Hölder. The revised work was translated by Robert B. Dean and Julius J. Brandstatter and published in two volumes as *Calculus of variations and partial differential equations of the first order* by Holden-Day in 1965–66. The Chelsea second edition combines and revises the 1967 edition.

[4] G. A. Jehle. 1991. *Advanced microeconomic theory*. Englewood Cliffs, New Jersey: Prentice-Hall.

[5] W. Karush. 1939. Minima of functions of several variables with inequalities as side conditions. Master's thesis, Department of Mathematics, University of Chicago.

[6] H. W. Kuhn. 1982. Nonlinear programming: A historical view. *ACM SIGMAP Bulletin* 31:6–18. Reprinted from SIAM–AMS Proceedings, volume IX, pp. 1–26.

[7] H. W. Kuhn and A. W. Tucker. 1951. Nonlinear programming. In J. Neyman, ed., *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability II, Part I*, pages 481–492. Berkeley: University of California Press. Reprinted in [8, Chapter 1, pp. 3–14]. http://projecteuclid.org/euclid.bsmsp/1200500249

[8] P. Newman, ed. 1968. *Readings in mathematical economics I: Value theory*. Baltimore: Johns Hopkins Press.

[9] J. P. Quirk. 1976. *Intermediate microeconomics: Mathematical notes*. Chicago: Science Research Associates.

[10] R. W. Shephard. 1981. *Cost and production functions*. Number 194 in Lecture Notes in Economics and Mathematical Systems. Berlin: Springer–Verlag. Reprint of the 1953 edition published by Princeton University Press.