Optimization, Day 2

Inequality constraints

$$\max_{x,y} U(x,y) \quad \text{s.t.} \quad P_x X + P_y Y = W$$
$$x, y \geq 0$$

Kuhn-Tucker conditions

$$U_x \leq \lambda P_x, \quad x \geq 0, \quad \overbrace{x \cdot (U_x - \lambda P_x) = 0}^{\text{complementary slackness}}$$
$$U_y \leq \lambda P_y, \; y \geq 0, \; \text{C.S.}$$
$$P_x X + P_y Y = W$$

"either the condition holds with equality, or $x = 0$"

if $x = 0$, the non-negativity constraint is **binding**, so the solution will be on the boundary, so the usual calculus conditions do not apply.

interior solution

corner solution

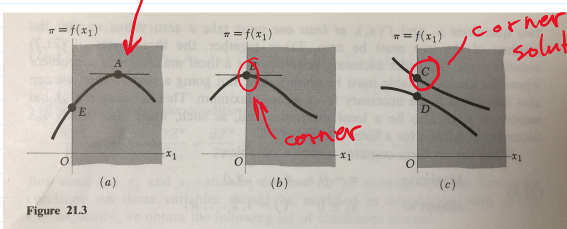
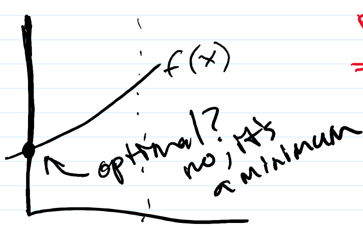


Figure 21.3

so $f'(A) = 0$ and $f'(B) = 0$

$x_1 \geq 0$
at C, the optimal point, $x_1 = 0$
$\Rightarrow$ slope < 0



if $f'(0) > 0$, then just increase x a little bit to get a better point

---

More general case

$$\max_{x,y} F(x,y) \quad \text{s.t.} \quad g(x,y) \leq c, \quad x, y \geq 0$$

FONC's

$$F_x \leq \lambda g_x, \quad x \geq 0, \quad x \cdot (F_x - \lambda g_x) = 0$$
$$F_y \leq \lambda g_y, \quad y \geq 0, \quad y \cdot (F_y - \lambda g_y) = 0$$
$$\underbrace{g(x,y) \leq c}_{\text{the constraint}}, \quad \lambda \geq 0, \quad \underbrace{\lambda \cdot (c - g(x,y))}_{\text{complementary slackness}} = 0$$

$\underbrace{g(x,y) \le c}_{\text{the constraint}}$, $\lambda \ge 0$, $\underbrace{\lambda \cdot (c - g(x,y))}_{\text{complementary slackness}} = 0$

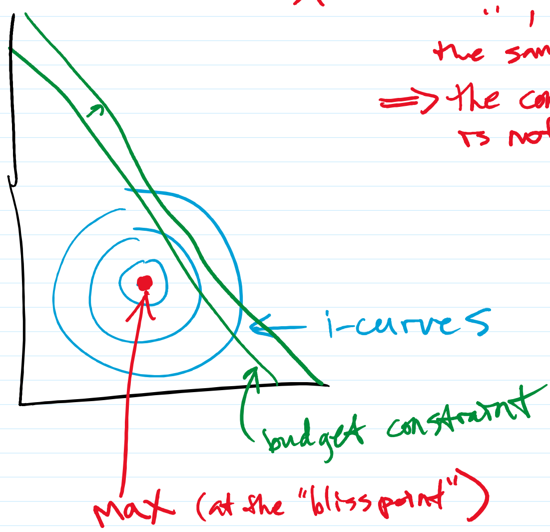
how does the value function change as you relax the constraint a little bit

complementary slackness
→ either $\lambda = 0$ (optimum on interior of constraint set)

or the constraint is binding at the optimum

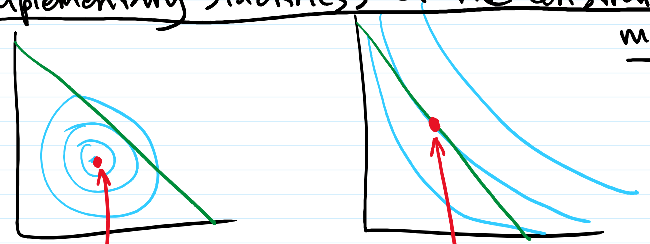
$\lambda > 0$ when you relax the constraint, you get a better optimum

$\lambda = 0$ " " " " " " the same optimum
⇒ the constraint is not binding



shifting BC out doesn't change your choice
$\Rightarrow \lambda = 0$

Complementary Slackness of the constraint and multiplier



$\lambda = 0$
$c - g(x) > 0$
i.e. $P_x x + P_y Y < W$

$\lambda > 0$
$P_x X + P_y Y = W$

General case (Kuhn-Tucker Conditions)

$$\max_{x_1, \dots, x_n} F(\vec{x}) \quad \text{s.t.} \quad c_i - g_i(\vec{x}) \ge 0 \quad i = 1, \dots, m$$

$$x_1, \dots, x_n \ge 0$$

FOC's
$$\frac{\partial F}{\partial x_1} \le \lambda_1 \frac{\partial g_1}{\partial x_1} + \lambda_2 \frac{\partial g_2}{\partial x_1} + \dots + \lambda_m \frac{\partial g_m}{\partial x_1}, \quad x_1 \ge 0, \quad \underbrace{x_1 \cdot (F_{x_1} - \vec{\lambda} \cdot \nabla g_{x_1}) = 0}_{\text{complementary slackness}}$$

$\vdots$

$$\frac{\partial F}{\partial x_n} \le \lambda_1 \frac{\partial g_1}{\partial x_n} + \dots + \lambda_m \frac{\partial g_m}{\partial x_n}, \quad x_n \ge 0, \quad x_n \cdot (F_{x_n} - \vec{\lambda} \nabla g_{x_n}) = 0$$

slackness

$$\frac{\partial F}{\partial x_n} \leq \lambda_1 \frac{\partial g_1}{\partial x_n} + \cdots + \lambda_m \frac{\partial g_m}{\partial x_n}, \quad x_n \geq 0, \quad x_n \cdot (F_{x_n} - \vec{\lambda} \nabla g_{x_n}) = 0$$

$$c_i - g_i(x) \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i \cdot (c_i - g_i(x)) = 0 \quad i = 1, \ldots, m$$

Necessary vs. Sufficient Conditions

$\nabla f = \lambda \nabla g$
Kuhn-Tucker Cond'ns] Necessary Conditions for a maximum → "First order conditions" → first order derivatives

$f''(x^*) < 0$
$H(x^*)$ negative define] sufficient conditions for a local max → "second order conditions" → 2nd order partial derivatives

Generally:

1) Solve the FOC's
2) Check the SOC's (or invoke an assumption that implies SOC's, like concavity)
3) If there are multiple candidates, evaluate each one and compare to find the global max.

Aside: Lagrangian

$$\mathcal{L} = F(x) + \lambda(c_1 - g_1(x)) + \mu(c_2 - g_2(x))$$
$$\mathcal{L} = F(x) - \lambda(g(x) - c_1) - \mu(g_2(x) - c_2)$$
same

The point: $\frac{\partial F}{\partial x_1} \leq \lambda g_{x_1} + \mu g_{x_1}$

Occasionally bad people write $\mathcal{L} = F(x) - \lambda(c_1 - g_1(x))$

then you need to change FOC's, $\lambda \leq 0$

Don't do this

K-T Thm, constraint qual
Lagrange Mult Thm, implicit fn thm
Concave Prog.
K-T sufficiency
Max Thm
Envelope Thm

Kuhn-Tucker theorem

Kuhn–Tucker Theorem: Suppose x is an n-dimensional vector, c an m-dimensional vector, F a function taking scalar values, G a function taking m-dimensional vector values. Define

$$L(x,\lambda) = F(x) + \lambda\,[\,c - G(x)\,], \tag{3.4}$$

where λ is an m-dimensional row vector. Suppose $\bar{x}$ maximizes $F(x)$ subject to $G(x) \le c$ and $x \ge 0$, and the constraint qualification holds, namely the submatrix of $G_x(\bar{x})$ formed by taking those rows i for which $G^i(\bar{x}) = c_i$ has the maximum possible rank. Then there is a value of λ such that

$$L_x(\bar{x},\lambda) \le 0,\ \bar{x} \ge 0, \quad \text{with complementary slackness,} \tag{3.7}$$

and

$$L_\lambda(\bar{x},\lambda) \ge 0,\ \lambda \ge 0, \quad \text{with complementary slackness.} \tag{3.10}$$

(Dixit, p181)

"If constraint qualification holds, then Kuhn-Tucker conditions are necessary for a local max."

Constraint Qual: Jacobian of the constraint functions has full row rank.

→ there are other types of constraint qualification, too.

→ see Simon+Blume, Chiang

Kuhn-Tucker Sufficiency (concave programming)

For the *maximization* problem, Kuhn and Tucker offer the following statement of sufficient conditions (sufficiency theorem):

Given the nonlinear program

Maximize $\quad \pi = f(x)$
subject to $\quad g^i(x) \le r_i \qquad (i = 1, 2, \ldots, m)$
and $\quad x \ge 0$

if the following conditions are satisfied:

(*a*) the objective function $f(x)$ is differentiable and *concave* in the nonnegative orthant *

(*b*) each constraint function $g^i(x)$ is differentiable and *convex* in the nonnegative orthant * → Arrow-Enthoven show this can also be quasiconvex

(*c*) the point $\bar{x}$ satisfies the Kuhn-Tucker maximum conditions *

then $\bar{x}$ gives a global maximum of $\pi = f(x)$.

If F is concave and g is convex then the K-T cond'ns are necessary and sufficient for a global max.

Implication: assume your objective function is concave whenever reasonable.

Arrow-Enthoven Sufficiency (quasi-concave programming)

Kenneth Arrow – GOAT

The Arrow-Enthoven Sufficiency Theorem

The theorem is as follows:

> Given the nonlinear program
>
> Maximize $\pi = f(x)$
>
> subject to $g^i(x) \le r_i \qquad (i = 1, 2, \ldots, m)$
>
> and $x \ge 0$
>
> if the following conditions are satisfied:
>
> (*a*) the objective function $f(x)$ is differentiable and *quasiconcave* in the nonnegative orthant
>
> (*b*) each constraint function $g^i(x)$ is differentiable and *quasiconvex* in the nonnegative orthant
>
> (*c*) the point $\bar{x}$ satisfies the Kuhn-Tucker maximum conditions
>
> (*d*) any *one* of the following is satisfied:
>
> (*d-i*) $f_j(\bar{x}) < 0$ for at least one variable x_j
>
> (*d-ii*) $f_j(\bar{x}) > 0$ for some variable x_j that can take on a positive value without violating the constraints
>
> (*d-iii*) the n derivatives $f_j(\bar{x})$ are not all zero, and the function $f(x)$ is twice differentiable in the neighborhood of $\bar{x}$ [i.e., all the second-order partial derivatives of $f(x)$ exist at $\bar{x}$]
>
> (*d-iv*) the function $f(x)$ is concave
>
> then $\bar{x}$ gives a global maximum of $\pi = f(x)$.

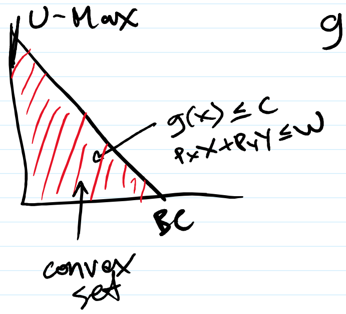
(d-iv) implies that concave programming really just requires that the constraint set is convex. It doesn't require $g(x)$ is a convex function, just quasiconvex.

Quasiconvex: lower contour set is convex

i.e.

$g(x) \le C$ is a convex set

U-Max



Maximum Theorem

Gives conditions under which a value function is a continuous fn of the parameters.

> **Theorem M.K.6:** (*Theorem of the Maximum*) Suppose that the constraint correspondence $C: Q \to \mathbb{R}^N$ is continuous (see Section M.H) and that $f(\cdot)$ is a continuous function. Then the maximizer correspondence $x: Q \to \mathbb{R}^N$ is upper hemicontinuous and the value function $v: Q \to \mathbb{R}$ is continuous.

(MWG Appendix)

For example:

$$\max_{x,y} U(x,y) \quad \text{s.t.} \quad P_x X + P_y y \le W$$

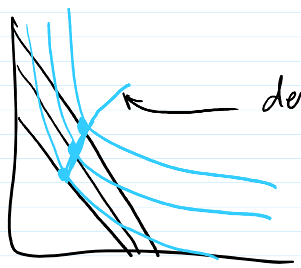
B is a correspondence of the parameters (p_x, p_y, w)

$$B : \underbrace{\mathbb{R}^3}_{\text{parameter space}} \rightrightarrows \underbrace{\mathbb{R}^2}_{\text{the consumption space}}$$

Can check if it is
UHC, LHC, or both (continuous)

If continuous, then apply the maximum theorem.

APEC 8004



demand correspondence $x^d(p, w)$

- if budget set continuous then this is upper hemicontinuous.

→ very useful for using fixed point theorems.

$V(p, w)$ – value fn (indirect utility fn) is also continuous.

AND if the maximizer correspondence (eg. demand correspondence) is single valued, then it is a continuous fn.

⇒ assuming strict convexity (eg. of preferences) implies that you get continuous, single-valued demand functions.

Envelope Theorem

Theorem M.L.1: (*Envelope Theorem*) Consider the value function $v(q)$ for the problem (M.L.1). Assume that it is differentiable at $\bar{q} \in \mathbb{R}^S$ and that $(\lambda_1, \ldots, \lambda_M)$ are values of the Lagrange multipliers associated with the maximizer solution $x(\bar{q})$ at $\bar{q}$. Then[29]

$$\frac{\partial v(\bar{q})}{\partial q_s} = \underbrace{\frac{\partial f(x(\bar{q}); \bar{q})}{\partial q_s}}_{\text{direct effect in unconstrained case}} - \underbrace{\sum_{m=1}^{M} \lambda_m \frac{\partial g_m(x(\bar{q}); \bar{q})}{\partial q_s}}_{\text{effect on constraint functions}} \quad \text{for } s = 1, \ldots, S, \qquad \text{(M.L.4)}$$

or, in matrix notation,

$$\nabla v(\bar{q}) = \nabla_q f(x(\bar{q}); \bar{q}) - \sum_{m=1}^{M} \lambda_m \nabla_q g_m(x(\bar{q}); \bar{q}). \qquad \text{(M.L.5)}$$

Proof: We proceed as in the case of a single variable and no constraints. Let $x(\cdot)$ be the maximizer function. Then $v(q) = f(x(q); q)$ for all q, and therefore, using the chain rule, we have

The derivative of the value fn wrt the parameters equals the derivative of the Lagrangian wrt those same parameters, evaluated at the optimal point.

Special case: no constraints.

Then $$\frac{dV(x;\alpha)}{d\alpha} = \frac{\partial F(x;\alpha)}{\partial \alpha}$$

where F is the objective function, V is the value fn, x are choice variables and α is a parameter.

Proof:

$$\frac{dV(\alpha)}{d\alpha} = \frac{dF(x^*;\alpha)}{d\alpha} = \underbrace{\sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \cdot \frac{\partial x_i^*}{\partial \alpha}}_{\text{indirect}} + \underbrace{\frac{\partial F}{\partial \alpha}}_{\text{direct}}$$

each $x_i^*(\alpha)$ is a function of α. $x_i^*(\alpha)$ are known as maximizers.

(note: $V(\alpha) = F(x^*(\alpha);\alpha)$)

And $\frac{\partial F}{\partial x_i} = 0$ at the max (F.O.C.)

so we get all the indirect effects go to 0. So...

$$\frac{dV(\alpha)}{d\alpha} = \underbrace{\frac{\partial F}{\partial \alpha}}_{\text{direct effect only}}$$

General Version: (w/ constraints)

$$\frac{dV(\alpha)}{d\alpha} = \frac{\partial F}{\partial \alpha} + \lambda \frac{\partial G}{\partial \alpha}$$

See MWG Appendix M.L.
for proof and visuals.

Lagrange Multiplier Theorem

Theorem 18.2 Let $f, h_1, \ldots, h_m$ be C^1 functions of n variables. Consider the problem of maximizing (or minimizing) $f(\mathbf{x})$ on the constraint set

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

Suppose that $\mathbf{x}^* \in C_{\mathbf{h}}$ and that $\mathbf{x}^*$ is a (local) max or min of f on $C_{\mathbf{h}}$. Suppose further that $\mathbf{x}^*$ satisfies condition NDCQ above. Then, there exist $\mu_1^*, \ldots, \mu_m^*$ such that $(x_1^*, \ldots, x_n^*, \mu_1^*, \ldots, \mu_m^*) \equiv (\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a critical point of the Lagrangian

$$L(\mathbf{x}, \boldsymbol{\mu}) \equiv f(\mathbf{x}) - \mu_1[h_1(\mathbf{x}) - a_1] - \mu_2[h_2(\mathbf{x}) - a_2] - \cdots - \mu_m[h_m(\mathbf{x}) - a_m].$$

In other words,

$$\boxed{\begin{aligned} \frac{\partial L}{\partial x_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, \ldots, \frac{\partial L}{\partial x_n}(\mathbf{x}^*, \boldsymbol{\mu}^*) &= 0, \\ \frac{\partial L}{\partial \mu_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, \ldots, \frac{\partial L}{\partial \mu_m}(\mathbf{x}^*, \boldsymbol{\mu}^*) &= 0. \end{aligned}} \quad (14)$$

(Simon + Blume)

$\mu_1 \cdots \mu_m$ Lagrange multipliers

In yet other words:

(*) $\nabla f = \mu_1 \nabla g_1 + \mu_2 \nabla g_2 + \cdots + \mu_m \nabla g_m$

this is necessary condition for a maximum.

NDCQ: nondegenerate constraint qualification

$J(h_1(x), h_2(x) \ldots, h_m(x))$ has full row rank.

$$J = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \frac{\partial h_m}{\partial x_2} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix} \leftarrow \text{one constraint}$$

$m \times n$ ↑ one variable

J needs $\text{rank}(J) = m$

i.e. the rows are linearly independent.

Aside: Implicit function theorem

One constraint

Theorem 15.2 (Implicit Function Theorem) Let $G(x_1, \ldots, x_k, y)$ be a C^1 function around the point $(x_1^*, \ldots, x_k^*, y^*)$. Suppose further that $(x_1^*, \ldots, x_k^*, y^*)$ satisfies

$$G(x_1^*, \ldots, x_k^*, y^*) = c$$

and that

$$\frac{\partial G}{\partial y}(x_1^*, \ldots, x_k^*, y^*) \neq 0.$$

If $G(x,y)$ defines

$$G(x_1^*, \ldots, x_k^*, y^*) = c$$

and that

$$\frac{\partial G}{\partial y}(x_1^*, \ldots, x_k^*, y^*) \neq 0.$$

Then, there is a C^1 function $y = y(x_1, \ldots, x_n)$ defined on an open ball B about $(x_1^*, \ldots, x_k^*)$ so that:

(a) $G(x_1, \ldots, x_k, y(x_1, \ldots, x_k)) = c$ for all $(x_1, \ldots, x_k) \in B$,
(b) $y^* = y(x_1^*, \ldots, x_k^*)$, and
(c) for each index i,

$$\frac{\partial y}{\partial x_i}(x_1^*, \ldots, x_k^*) = -\frac{\dfrac{\partial G}{\partial x_i}(x_1^*, \ldots, x_k^*, y^*)}{\dfrac{\partial G}{\partial y}(x_1^*, \ldots, x_k^*, y^*)}. \quad (15)$$

If $G(x,y)$ defines y implicitly as a function of x, then

$$\frac{dy}{dx} = \frac{-\partial G/\partial x}{\partial G/\partial y}$$

Theorem 15.7 Let $F_1, \ldots, F_m \colon \mathbf{R}^{m+n} \to \mathbf{R}^1$ be C^1 functions. Consider the system of equations

$$\begin{aligned} F_1(y_1, \ldots, y_m, x_1, \ldots, x_n) &= c_1 \\ \vdots \qquad & \quad \vdots \\ F_m(y_1, \ldots, y_m, x_1, \ldots, x_n) &= c_m \end{aligned} \quad (32)$$

m equations and m possible implicit functions.

as possibly defining $y_1, \ldots, y_m$ as implicit functions of $x_1, \ldots, x_n$. Suppose that $(\mathbf{y}^*, \mathbf{x}^*)$ is a solution of (32). If the determinant of the $m \times m$ matrix

$$\begin{pmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_m} \end{pmatrix} \equiv \frac{\partial(F_1, \ldots, F_h, \ldots, F_m)}{\partial(y_1, \ldots, y_h, \ldots, y_m)}$$

Note: If y, then $\frac{\partial F}{\partial y} \neq$

m of the $(m+n)$ are endogenous (can be expressed as functions of the other n exogenous variables).

evaluated at $(\mathbf{y}^*, \mathbf{x}^*)$ is nonzero, then there exist C^1 functions

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_n) \\ &\ \ \vdots \qquad \vdots \\ y_m &= f_m(x_1, \ldots, x_n) \end{aligned} \quad (33)$$

defined on a ball B about $\mathbf{x}^*$ such that

$$\begin{aligned} F_1(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}), x_1, \ldots, x_n) &= c_1 \\ \vdots \qquad & \quad \vdots \\ F_m(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}), x_1, \ldots, x_n) &= c_m \end{aligned}$$

for all $\mathbf{x} = (x_1, \ldots, x_n)$ in B and

$$\begin{aligned} y_1^* &= f_1(x_1^*, \ldots, x_n^*) \\ \vdots \quad & \quad \vdots \\ y_m^* &= f_m(x_1^*, \ldots, x_n^*). \end{aligned}$$

Furthermore, one can compute $(\partial f_k/\partial x_h)(\mathbf{y}^*, \mathbf{x}^*) = (\partial y_k/\partial x_h)(\mathbf{y}^*, \mathbf{x}^*)$ by setting $dx_h = 1$ and $dx_j = 0$ for $j \neq h$ in (27) and solving the resulting system for dy_k. This can be accomplished:

linearization, 353

(a) by inverting the nonsingular matrix (28) to obtain the solution (30) or
(b) by applying Cramer's rule to (27) to obtain the solution (31).

Determinant of the Jacobian of the system w.r.t. the endogenous variables must be nonzero, i.e. it is nonsingular.

In the proof of the L-M theorem, we have $x_1, \ldots, x_n$ and m constraints where $m < n$.

$$\underbrace{x_1, \ldots, x_m}_{\text{endogenous}}, \underbrace{x_{m+1}, \ldots x_n}_{\text{exogenous}}$$

endogenous exogenous

Proof of L-M Theorem

$$J = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \\ \frac{\partial g_m}{\partial x_1} & & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

(f - obj. fn
g_i - constraint fn)

$(m+1) \times n$

Step 1) Show J has rank $< m+1$
We do this as a proof by contradiction.

Step 2) Show that this implies

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_m \nabla g_m$$

We have x^* which maximizes:

$$A \begin{cases} f(x) = c_0 \\ g_1(x) = c_1 \\ g_2(x) = c_2 \\ \vdots \\ g_m(x) = c_m \end{cases}$$

$c_0 \equiv f(x^*)$ the max value

J is the Jacobian of this system.

Suppose J has rank $= m+1$ so that all rows are linearly indep.

Then consider A as a system of implicit functions where

$x_1, \dots, x_m$ are endogenous

and $x_{m+1}, \dots, x_n$ and $c_0, c_1, \dots, c_m$

are exogenous.

Then there is an open ball around

$$\underbrace{x^*_{m+1}, \dots, x^*_n}_{\text{optimal value}}, c_0, \dots, c_m$$

such that the system is satisfied for all points in that ball (implicit fn thm).

In particular, consider a point in that ball, $x^{**}_{m+1}, \dots, x^{**}_n, c_0+\varepsilon, c_1, \dots, c_m$.

The implicit fn thm says A is still satisfied at this point. This implies the constraints are satisfied at x^{**}.

But $f(x^{**}) = c_0 + \varepsilon > c_0 = f(x^*)$

But this contradicts that $f(x^*)$ was a maximum.

So... J can't have rank of $m+1$.

This implies the rows of J are linearly dependent. So $\underbrace{\lambda_0 \nabla f + \lambda_1 \nabla g_1 + \cdots + \lambda_m \nabla g_m = 0}_{\text{homogeneous system}}$ where $\lambda_0, \dots, \lambda_m$ are not all 0.

→ nontrivial sol'ns where the cols are linearly dep.

(note: ∇g_1 etc. were rows in J but now they are columns, but this doesn't change the problem)

By the NDCQ, we know

By the NDCQ, we know
$\nabla g_1, \dots, \nabla g_m$ are all linearly indep.
So this $\Rightarrow \lambda_0 \neq 0$. Otherwise
we'd have $\lambda_1 \nabla g_1 + \cdots + \lambda_m \nabla g_m = 0$, which
contradicts linear indep. since a homogeneous
system w/ linearly indep. cols has only the trivial
solution.

$$\nabla f = \underbrace{\frac{\lambda_1}{\lambda_0}}_{\text{multiplier}} \nabla g_1 + \cdots \frac{\lambda_m}{\lambda_0} \nabla g_m$$

Q.E.D.

See Simon + Blume p. 478.