19. Constrained Optimization II

11/5/20

NB: Problems 4 and 7 from Chapter 17 and problems 5, 9, 11, and 15 from Chapter 18 are due on Thursday, November 12.

We continue our investigation of constrained optimization, including some the ideas surrounding the Kuhn-Tucker theory. We focus on four main areas.

1. Dependence of the optimum and optimal values on model parameters, including the Envelope Theorem.
2. Second order conditions using bordered Hessians.
3. Smoothness of Optimal Functions.
4. Constraint qualification: Necessity and alternatives.

19.1 Interpreting the Multiplier

We've used multipliers to solve optimization problems, but we haven't stopped to ask: What does the multiplier mean?

To answer that, let's go back to the basic relation involving the multiplier and examine it in a simple model. We have an objective function f and a single constraint function h. We solve the following maximization problem.

$$\begin{aligned} &\max_{x} f(x, y) \\ &\text{ s.t. } h(x, y) = c \end{aligned} \tag{19.1.1}$$

The optimality condition defines the multiplier μ.

$$\nabla f = \mu \nabla h.$$

Even in the simplest cases where it's useful to use a multiplier, we must have at least two goods. So we have a vector equation. Let's focus on the first component.

$$\frac{\partial f}{\partial x} = \mu \frac{\partial h}{\partial x} \quad \text{or} \quad \mu = \frac{\partial f}{\partial x} \bigg/ \frac{\partial h}{\partial x}$$

The units of μ are units of f per unit of h. This happens regardless of which variable, x or y, is used to define μ. That suggests the multiplier is the marginal value of whatever it is that is being constrained, whatever h is. In this model we can all call it c. That suggests we calculate the marginal value of c and see if it is in fact the multiplier.

19.2 A Multiplier Theorem

We return to the maximization problem

$$\begin{aligned} &\max_{(x,y)} f(x,y) \\ &\text{ s.t. } h(x,y) = c \end{aligned} \tag{19.1.1}$$

and regard c is an adjustable parameter. The solutions depend on it, so we write $\big(x^*(c), y^*(c)\big)$. The optimal value of f also depends on c. It is $f\big(x^*(c), y^*(c)\big)$.

What we want to know is whether

$$\frac{d}{dc}\Big[f\big(x^*(c), y^*(c)\big)\Big] = \mu.$$

In fact, it does.

Theorem 19.2.1. *Let f and h be $\mathcal{C}^1$ functions on $U \subset \mathbb{R}^2$. Suppose the optimal solution to equation (19.1.1), $\big(x^*(c), y^*(c)\big)$, is a continuously differentiable function of c, as is the corresponding multiplier $\mu^*(c)$ and that NDCQ holds at $\big(x^*(c), y^*(c), \mu^*(c)\big)$. Then*

$$\mu^*(c) = \frac{d}{dc}\Big[f\big(x^*(c), y^*(c)\big)\Big].$$

19.3 Calculating the Multiplier

Theorem 19.2.1. *Let* f *and* h *be* $\mathcal{C}^1$ *functions on* $U \subset \mathbb{R}^2$*. Suppose the optimal solution to equation (19.1.1),* $\big(x^*(c), y^*(c)\big)$*, is a continuously differentiable function of* c*, as is the corresponding multiplier* $\mu^*(c)$ *and that NDCQ holds at* $\big(x^*(c), y^*(c), \mu^*(c)\big)$*. Then*

$$\mu^*(c) = \frac{d}{dc}\Big[f\big(x^*(c), y^*(c)\big)\Big].$$

Proof. The Lagrangian is $\mathcal{L} = f(x,y) - \mu\big(h(x,y) - c\big)$ The first order conditions are

$$0 = \frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} - \mu^* \frac{\partial h}{\partial x} \qquad 0 = \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial f}{\partial y} - \mu^* \frac{\partial h}{\partial y}$$

so

$$\frac{\partial f}{\partial x} = \mu^* \frac{\partial h}{\partial x} \quad \text{and} \quad \frac{\partial f}{\partial y} = \mu^* \frac{\partial h}{\partial y} \tag{19.3.2}$$

with everything evaluated at $(x, y, \mu) = \big(x^*(c), y^*(c), \mu^*(c)\big)$. Differentiating the constraint $h\big(x^*(c), y^*(c)\big) = c$, we obtain

$$\frac{\partial h}{\partial x}\frac{dx^*}{dc} + \frac{\partial h}{\partial y}\frac{dy^*}{dc} = 1. \tag{19.3.3}$$

Applying the Chain Rule and using equations (19.3.2) and (19.3.3) we find

$$\begin{aligned}
\frac{d}{dc}\Big[f\big(x^*(c), y^*(c)\big)\Big] &= \frac{\partial f}{\partial x}\frac{dx^*}{dc} + \frac{\partial f}{\partial y}\frac{dy^*}{dc} \\
&= \left(\mu^* \frac{\partial h}{\partial x}\right)\frac{dx^*}{dc} + \left(\mu^* \frac{\partial h}{\partial y}\right)\frac{dy^*}{dc} && \text{By (19.3.2)} \\
&= \mu^* \left(\frac{\partial h}{\partial x}\frac{dx^*}{dc} + \frac{\partial h}{\partial y}\frac{dy^*}{dc}\right) \\
&= \mu^*. && \text{By (19.3.3)}
\end{aligned}$$

∎

Theorem 19.2.1 is a special case of the Envelope Theorem, which we will encounter later. All varieties of the theorem have very similar proofs. We use the Chain Rule to calculate the derivative of the optimal value $f\big(x^*(c), y^*(c)\big)$, which we then rewrote twice: Once using the first order conditions (19.3.2), and a second time using the results (19.3.3) from differentiating the constraint.

19.4 Marginal Utility of Income and Marginal Cost

Let's apply Theorem 19.2.1 to the simple consumer's problem we examined in the previous chapter.

$$\begin{aligned} v(\mathbf{p}, m) = \max_{\boldsymbol{x}}\, & u(x_1, x_2) \\ \text{s.t. } & p_1 x_1 + p_2 x_2 = m. \end{aligned}$$

We treat m as the parameter. Provided the solutions are differentiable, (NDCQ holds automatically), we can apply Theorem 19.2.1 to find

$$\frac{\partial v}{\partial m}(\mathbf{p}, m) = \mu^*(m).$$

The multiplier belonging to the budget constraint is the marginal (indirect) utility of income.

We can also consider a simplified version of the firm's cost minimization problem.

$$\begin{aligned} c(\boldsymbol{w}, q) = \min_{\boldsymbol{z}}\, & \boldsymbol{w}\cdot\boldsymbol{z} \\ \text{s.t. } & f(\boldsymbol{z}) = q, . \end{aligned}$$

We assume $Df \gg \mathbf{0}$, insuring that NDCQ holds. If the solutions are $\mathcal{C}^1$, we have

$$\frac{\partial c}{\partial q}(\boldsymbol{w}, q) = \mu^*(q).$$

The multiplier here is the marginal cost.

Of course, we have to be careful drawing strong conclusions from these simple versions of the utility maximization and cost minimization problems. The details left out can be important. We will generalize Theorem 19.2.1 to apply to the full consumer's utility maximization problem and firm's cost minimization problem. That version can be used to establish these results for the standard consumer's utility maximization problem and firm's cost minimization problem.

19.5 Multiple Multipliers

Suppose we have many variables and multiple constraints, with multiple multipliers. We consider the problem

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}) \\ &\text{s.t. } h_i(\mathbf{x}) = c_i, \text{ for } i = 1, \ldots, \ell \end{aligned} \tag{19.5.4}$$

A version of Theorem 19.2.1 applies here too. Each multiplier corresponds to a different derivative of $f\big(\mathbf{x}^*(\mathbf{c})\big)$.

Theorem 19.5.1. *Let* $U \subset \mathbb{R}^m$ *and suppose* $f\colon U \to \mathbb{R}$ *and* $\mathbf{h}\colon U \to \mathbb{R}^\ell$ *are* $\mathcal{C}^1$ *functions. Suppose the optimal solution to equation (19.5.4),* $\mathbf{x}^*(\mathbf{c})$*, is* $\mathcal{C}^1$ *in* $\mathbf{c}$*, as is the corresponding vector of multipliers* $\boldsymbol{\mu}^*(\mathbf{c})$*, and finally that NDCQ holds at* $\big(\mathbf{x}^*(\mathbf{c}), \boldsymbol{\mu}^*(\mathbf{c})\big)$*. Then*

$$\boldsymbol{\mu}^*(\mathbf{c})^{\mathsf{T}} = D_{\mathbf{c}}\Big[f\big(\mathbf{x}^*(\mathbf{c})\big)\Big]$$

Proof. The first order conditions are

$$D_{\mathbf{x}} f = \boldsymbol{\mu}^{*\mathsf{T}} D_{\mathbf{x}} \mathbf{h} \tag{19.5.5}$$

where everything is evaluated at the optimum. We also differentiate the constraints $\mathbf{h}\big(\mathbf{x}^*(\mathbf{c})\big) = \mathbf{c}$, yielding

$$(D_{\mathbf{x}}\mathbf{h})(D_{\mathbf{c}}\mathbf{x}^*) = \mathbf{I}_{\ell\times\ell} \tag{19.5.6}$$

Use the Chain Rule to differentiate $f\big(\mathbf{x}^*(c), c\big)$ and then substitute equations (19.5.5) and (19.5.6) to complete the proof:

$$\begin{aligned} D_{\mathbf{c}}\Big[f\big(\mathbf{x}^*(\mathbf{c})\big)\Big] &= (D_{\mathbf{x}} f)(D_{\mathbf{c}}\mathbf{x}^*) && \text{Chain Rule} \\ &= \boldsymbol{\mu}^*(\mathbf{c})^{\mathsf{T}}(D_{\mathbf{x}}\mathbf{h})(D_{\mathbf{c}}x^*) && \text{By equation (19.5.5)} \\ &= \boldsymbol{\mu}^*(\mathbf{c})^{\mathsf{T}}\mathbf{I}_{\ell\times\ell} && \text{By equation (19.5.6)} \\ &= \boldsymbol{\mu}^*(\mathbf{c})^{\mathsf{T}} \end{aligned}$$

∎

19.6 Multipliers and Mixed Constraints

We move on to the general case of mixed constraints. This means we have an optimization problem of this form.

$$\begin{aligned} &\max_{\mathbf{x}}\ f(\mathbf{x}) \\ &\text{s.t. } g_i(\mathbf{x}) \le b_i, \text{ for } i = 1, \dots, k \\ &\qquad h_i(\mathbf{x}) = c_i, \text{ for } i = 1, \dots, \ell \end{aligned} \tag{19.6.7}$$

It's easy to see that if a constraint $g_h(\mathbf{x}^*) \le b_h$ does not bind at $\mathbf{x}^*$, that it will still not bind if we change b_h slightly. It follows that the maximized value, $f\big(\mathbf{x}^*(\mathbf{b},\mathbf{c})\big)$, does not change either, and that its b_h derivative must be zero. Since constraint h does not bind, the corresponding multiplier is zero by complementary slackness. The multiplier and derivative are the same.

The proof of the following theorem is similar to the previous results.

Theorem 19.6.1. *Let $U \subset \mathbb{R}^m$ and suppose $f\colon U \to \mathbb{R}$, $\mathbf{g}\colon U \to \mathbb{R}^k$, and $\mathbf{h}\colon U \to \mathbb{R}^\ell$ are $\mathcal{C}^1$ functions. Suppose the optimal solution to equation (19.6.7), $\mathbf{x}^*(\mathbf{b},\mathbf{c})$, is $\mathcal{C}^1$ in $(\mathbf{b},\mathbf{c})$, as are the corresponding vectors of multipliers $\boldsymbol{\mu}^*(\mathbf{b},\mathbf{c})$ and $\boldsymbol{\lambda}^*(\mathbf{b},\mathbf{c})$.*

Let $\hat{\mathbf{g}}$ be the vector of the $\hat{k}$ binding constraints at $\mathbf{x}^*$ and define

$$\mathbf{G}(\mathbf{x}) = \begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}.$$

Suppose $\operatorname{rank} D\mathbf{G}(\mathbf{x}^*) = \hat{k} + \ell < m$ *holds (NDCQ).*

Then

$$\begin{aligned} \boldsymbol{\lambda}^*(\mathbf{b},\mathbf{c})^{\mathsf{T}} &= D_{\mathbf{b}}\Big[f\big(\mathbf{x}^*(\mathbf{b},\mathbf{c})\big)\Big] \\ \boldsymbol{\mu}^*(\mathbf{b},\mathbf{c})^{\mathsf{T}} &= D_{\mathbf{c}}\Big[f\big(\mathbf{x}^*(\mathbf{b},\mathbf{c})\big)\Big]. \end{aligned}$$

Marginal Values. In all of the above theorems,

$$\frac{\partial f}{\partial b_i}\big(\mathbf{x}^*(\mathbf{b},\mathbf{c})\big) = \lambda_i \quad \text{and} \quad \frac{\partial f}{\partial c_j}\big(\mathbf{x}^*(\mathbf{b},\mathbf{c})\big) = \mu_j$$

are the marginal values of the maximized function when either b_i or c_j is increased.

19.7 Envelope Theorems

In the above theorem, we've been able to calculate the marginal value of relaxing any of the constraints. However, we would like to calculate marginal values for other parameters of the model, such as prices. For that, we need a generalization known as the Envelope Theorem.

Let $\mathbf{a} \in \mathbb{R}^n$ be a vector of parameters. Suppose the objective depends on both $\mathbf{x}$ and $\mathbf{a}$, with $f(\mathbf{x}, \mathbf{a}) \in \mathcal{C}^1$ and suppose we can also write $\mathbf{h}(\mathbf{x}, \mathbf{a}) \in \mathcal{C}^1$ where $\mathbf{h}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$ is the vector of equality constraints. Theorem 19.5.1 is just the special case with $\mathbf{h}(\mathbf{x}, \mathbf{a}) = \mathbf{h}(\mathbf{x}) - \mathbf{a}$.

The Lagrangian is

$$\mathcal{L} = f(\mathbf{x}, \mathbf{a}) - \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}, \mathbf{a})$$

We now employ the same procedure as before. The first order conditions are

$$0 = D_{\mathbf{x}}\mathcal{L} = D_{\mathbf{x}} f - (\boldsymbol{\mu}^*)^T D_{\mathbf{x}}\mathbf{h},$$

so

$$D_{\mathbf{x}} f = (\boldsymbol{\mu}^*)^T D_{\mathbf{x}}\mathbf{h}. \tag{19.7.8}$$

We differentiate the constraint, which now has an extra term

$$D_{\mathbf{a}}\Big[\mathbf{h}\big(\mathbf{x}^*(\mathbf{a}), \mathbf{a}\big)\Big] = (D_{\mathbf{x}}\mathbf{h})\,(D_{\mathbf{a}}\mathbf{x}^*) + D_{\mathbf{a}}\mathbf{h} = \mathbf{0}$$

so

$$(D_{\mathbf{x}}\mathbf{h})\,(D_{\mathbf{a}}\mathbf{x}^*) = -D_{\mathbf{a}}\mathbf{h} \tag{19.7.9}$$

Then we wrap it all together using the Chain Rule and substitute equations (19.7.8) and (19.7.9)

$$\begin{aligned}
D_{\mathbf{a}}\Big[f\big(\mathbf{x}^*(\mathbf{a}), \mathbf{a}\big)\Big] &= (D_{\mathbf{x}} f)\,(D_{\mathbf{a}}\mathbf{x}^*) + D_{\mathbf{a}} f \\
&= \big((\boldsymbol{\mu}^*)^T D_{\mathbf{x}}\mathbf{h}\big) D_{\mathbf{a}}\mathbf{x}^* + D_{\mathbf{a}} f && \text{By equation (19.7.8)} \\
&= -(\boldsymbol{\mu}^*)^T D_{\mathbf{a}}\mathbf{h} + D_{\mathbf{a}} f && \text{By equation (19.7.9)} \\
&= D_{\mathbf{a}}\Big[f - \boldsymbol{\mu}^{*T}\mathbf{h}\Big] \\
&= D_{\mathbf{a}}\mathcal{L}\big(\mathbf{x}^*(\mathbf{a}), \boldsymbol{\mu}^*(\mathbf{a}), \mathbf{a}\big) && \text{Definition of Lagrangian}
\end{aligned}$$

19.8 Statement of Envelope Theorem

We now state an Envelope Theorem based on the above calculations. We just need to add appropriate hypotheses so that we can carry out the preceding calculations.

Consider the problem

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}) \\ &\text{s.t. } h_j(\mathbf{x}, \mathbf{a}) = 0, \text{ for } j = 1, \dots, \ell \end{aligned} \tag{19.8.10}$$

Theorem 19.8.1. *Let $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ and suppose $f\colon U \times V \to \mathbb{R}$ and $\mathbf{h}\colon U \times V \to \mathbb{R}^\ell$ are $\mathcal{C}^1$ functions. Suppose $\mathbf{x}^*(\mathbf{a})$ solves equation (19.8.10) for any $\mathbf{a} \in V$. Let*

$$\mathcal{L} = f(\mathbf{x}, \mathbf{a}) - \boldsymbol{\mu}^{\mathsf{T}}\mathbf{h}(\mathbf{x}, \mathbf{a})$$

be the Lagrangian for this problem. Suppose that optimal solution $\mathbf{x}^*(\mathbf{a})$ and the corresponding Lagrange multipliers $\boldsymbol{\mu}^*(\mathbf{a})$ are $\mathcal{C}^1$ functions of $\mathbf{a}$, and that NDCQ holds. Then

$$D_{\mathbf{a}}\Big[f\big(\mathbf{x}^*(\mathbf{a}), \mathbf{a}\big)\Big] = D_{\mathbf{a}}\mathcal{L}\big(\mathbf{x}^*(\mathbf{a}), \boldsymbol{\mu}^*(\mathbf{a}), \mathbf{a}\big).$$

19.9 Some Applications of the Envelope Theorem

Consider a simple consumer's problem.

▶ Example 19.9.1: Roy's Identity. The indirect utility function $v(\mathbf{p}, m)$ defined by

$$\begin{aligned} v(\mathbf{p}, m) = \max_{\boldsymbol{x}}\ & u(\boldsymbol{x}) \\ \text{s.t. } & \mathbf{p}\cdot\boldsymbol{x} = m \end{aligned}$$

where $u\colon \mathbb{R}^m_+ \to \mathbb{R}$ is a utility function, the prices obey $\mathbf{p} \gg \mathbf{0}$ and income is positive, $m > 0$.

We set $\mathbf{a} = \mathbf{p}$. Then

$$\frac{\partial v}{\partial p_i} = \big(D_{\mathbf{p}}\mathcal{L}\big)_i = \frac{\partial u}{\partial p_i} - \mu^*(\mathbf{a})x_i(\mathbf{p}, m) = -\mu^*(\mathbf{a})x_i(\mathbf{p}, m).$$

We can use the fact that $\partial v/\partial m = \mu^*$ to establish *Roy's Identity*:

$$x_i(\mathbf{p}, m) = -\frac{\partial v/\partial p_i}{\partial v/\partial m}.$$

◀

Another example involves the cost function.

▶ Example 19.9.2: Shephard's Lemma. The cost function $c(\boldsymbol{w}, q)$ is defined by

$$\begin{aligned} c(\boldsymbol{w}, q) = \min_{\boldsymbol{z}}\ & \boldsymbol{w}\cdot\boldsymbol{z} \\ \text{s.t. } & f(\boldsymbol{z}) = q, . \end{aligned}$$

where $f\colon \mathbb{R}^m_+ \to \mathbb{R}$ is a production function and $\boldsymbol{w} \gg \mathbf{0}$ is the vector of factor prices.

Since $\boldsymbol{w}$ doesn't appear in the constraints, $D_{\boldsymbol{w}}c(\boldsymbol{w}, q) = D_{\boldsymbol{w}}(\boldsymbol{w}\cdot\boldsymbol{z}) = \boldsymbol{z}(\boldsymbol{w}, q)^{\mathsf{T}}$, the *conditional factor demands*. This result is known as *Shephard's Lemma*. ◀

We can also apply this to the very similar expenditure function.

▶ Example 19.9.3: Shephard-McKenzie Lemma. The expenditure function solves

$$\begin{aligned} e(\mathbf{p}, \bar{u}) = \min_{\boldsymbol{x}}\ & \mathbf{p}\cdot\boldsymbol{x} \\ \text{s.t. } & u(\boldsymbol{x}) = \bar{u}, . \end{aligned}$$

where $D_{\mathbf{p}}e(\mathbf{p}, \bar{u}) = D_{\mathbf{p}}(\mathbf{p}\cdot\boldsymbol{x}) = \mathbf{h}(\mathbf{p}, \bar{u})^{\mathsf{T}}$, the *Hicksian* or *compensated demands*. This result is sometimes called the *Shephard-McKenzie Lemma*. ◀

19.10 Second Order Sufficient Conditions: Maximum

As in the unconstrained case, we can use the second derivative to show that a critical point is maximal. For the sufficient conditions, we need to show that $D^2_{\mathbf{x}}\mathcal{L}$ is negative definite, at least for those vectors that are in the tangent space of the constraint set.

We can apply the bordered Hessian test to see if $D^2_{\mathbf{x}}\mathcal{L}$ is negative definite on $T_{\mathbf{x}^*}M = \ker D_{\mathbf{x}^*}\mathbf{h}$. Define the bordered Hessian of the Lagrangian by

$$\mathbf{B} = \begin{pmatrix} \mathbf{0}_{\ell\times\ell} & D\mathbf{h}(\mathbf{x}^*) \\ D\mathbf{h}(\mathbf{x}^*)^T & D^2_{\mathbf{x}}\mathcal{L} \end{pmatrix}.$$

where $\mathbf{0}_{\ell\times\ell}$ is the $\ell\times\ell$ zero matrix.

We can now state second order sufficient conditions for a constrained strict local maximum.

Theorem 19.10.1. *Let $U \subset \mathbb{R}^m$ and suppose $f\colon U \to \mathbb{R}$ and $\mathbf{h}\colon U \to \mathbb{R}^\ell$ are $\mathcal{C}^2$ functions. Suppose further that* $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$ *holds (NDCQ) and define the constraint set by $M = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$. The Lagrangian is*

$$\mathcal{L} = f(\mathbf{x}) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

Suppose

1. $\mathbf{x}^* \in M$.
2. *There are $\boldsymbol{\mu}^* \in \mathbb{R}^\ell$ with $D_{\mathbf{x}}\mathcal{L} = \mathbf{0}$ and $D_{\boldsymbol{\mu}}\mathcal{L} = \mathbf{0}$ at $(\mathbf{x}^*, \boldsymbol{\mu}^*)$.*
3. *The Hessian of the Lagrangian with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \boldsymbol{\mu}^*)$, $D^2_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\mu}^*)$, is negative definite on the tangent space $T_{\mathbf{x}^*}M = \ker D_{\mathbf{x}^*}\mathbf{h} = \{\mathbf{v} : \big(D_{\mathbf{x}^*}\mathbf{h}\big)\mathbf{v} = \mathbf{0}\}$.*

Then $\mathbf{x}^*$ is a strict local constrained maximum of f on M.

Moreover, the conclusion remains true if (3) is replaced by (3′).

3′. *The last $(m-\ell)$ leading principal minors of the bordered Hessian $\mathbf{B}$ alternate in sign and the determinant of $\mathbf{B}$ obeys $(-1)^{\ell+m}\det\mathbf{B} > 0$.*

It is arguably more natural to use $D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}$ rather than the bordered Hessian. Here

$$D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\mu}^*) = \begin{pmatrix} \mathbf{0}_{\ell\times\ell} & -D\mathbf{h}(\mathbf{x}^*) \\ -D\mathbf{h}(\mathbf{x}^*)^T & D^2_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\mu}^*) \end{pmatrix}$$

The extra minus signs do not affect whether the matrix is constrained positive definite or constrained negative definite. Compared to the bordered Hessian, the principal minors are unchanged because for each column that is multiplied by (-1), the corresponding row is also multiplied by (-1). Together, they multiply the determinant by $(-1)^2$, leaving every minor unchanged.

The bordered Hessian that appears in the proof of Theorem 19.10.1 is the natural bordered Hessian, $D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}$, not our usual bordered Hessian. Since they are equivalent in the sense that they both are negative definite on the same subspaces and yield the same conditions on minors, it doesn't really affect the statement of the theorem.

19.11 Proof of Theorem 19.10.1

Proof. We will focus on the Hessian of the Lagrangian with respect to μ and x: $(D^2_{(\mu,x)}\mathcal{L})(x^*,\mu^*)$. Consider the function

$$\phi(x,y) = \begin{pmatrix} -\mu^{*\mathsf{T}} & y^{\mathsf{T}} \end{pmatrix} \Big[\big(D^2_{(\mu^*,x)}\mathcal{L}\big)(x,\mu^*)\Big] \begin{pmatrix} -\mu^* \\ y \end{pmatrix}.$$

We restrict the second component to the $(m-1)$-sphere $S^{m-1} = \{y \in \mathbb{R}^m : \|y\| = 1\}$. Then ϕ is a continuous function of (x,y) defined on $U \times S^{m-1} \subset \mathbb{R}^{2m}$, as are all of its leading principal minors. The leading principal minors characterize negative definiteness of the natural bordered Hessian

$$\big(D^2_{(\mu,x)}\mathcal{L}\big)(x,\mu^*)$$

on the tangent space $T_{x^*}M = \ker D_{x^*}h$.

Due to the continuity of ϕ we can find an $\varepsilon_0 > 0$ and an open set V containing $T_{x^*}M \cap S^{m-1}$ with

$$\phi(x,y) < 0 \quad \text{for all } (x,y) \text{ in } B_{\varepsilon_0}(x^*) \times V.$$

Here x and y are not limited to M, although we will soon restrict x, but not y to M.

Take $\varepsilon < \varepsilon_0$ so that if $x \in B_\varepsilon(x^*) \cap M$ with $x \neq x^*$, $(x - x^*)/\|x - x^*\| \in V$. Then for any $x \in B_\varepsilon(x)$, Taylor's formula yields

$$f(x) = f(x^*) + \big[Df(x^*)\big](x - x^*) + \frac{1}{2}(x - x^*)^{\mathsf{T}}\big[D^2 f(y)\big](x - x^*).$$

For some $y \in \ell(x,x^*) \subset B_\varepsilon(x^*)$. It may be that $y \notin M$. Now $Df(x^*) = \mu^{*\mathsf{T}}Dh$ by the first order condition $D_x\mathcal{L} = 0$, so

$$f(x) = f(x^*) + \mu^{*\mathsf{T}}(Dh)(x - x^*) + \frac{1}{2}(x - x^*)^{\mathsf{T}} D^2 f(x)(x - x^*). \tag{19.11.11}$$

We calculate

$$\begin{aligned}
&\begin{pmatrix} -\mu^{*\mathsf{T}} & (x - x^*)^{\mathsf{T}} \end{pmatrix} \Big[\big(D^2_{(\mu^*,x)}\mathcal{L}\big)(y,\mu^*)\Big] \begin{pmatrix} -\mu^* \\ x - x^* \end{pmatrix}. \\
&= \begin{pmatrix} -\mu^{*\mathsf{T}} & (x - x^*)^{\mathsf{T}} \end{pmatrix} \begin{pmatrix} 0_{\ell\times\ell} & -Dh(x^*) \\ -Dh(x^*)^{\mathsf{T}} & D^2_x\mathcal{L}(x^*,\mu^*) \end{pmatrix} \begin{pmatrix} -\mu^* \\ x - x^* \end{pmatrix}. \\
&= \mu^{*\mathsf{T}}\big(Dh\big)(x - x^*) + (x - x^*)^{\mathsf{T}}\big(Dh\big)\mu^* + (x - x^*)^{\mathsf{T}}\big(D^2_y f\big)(x - x^*) \\
&= 2\mu^{*\mathsf{T}}\big(Dh\big)(x - x^*) + (x - x^*)^{\mathsf{T}}\big(D^2_y f\big)(x - x^*).
\end{aligned}$$

We can rewrite equation (19.11.11) as

$$f(x) = f(x^*) + \frac{1}{2}\begin{pmatrix} -\mu^{*\mathsf{T}} & (x - x^*)^{\mathsf{T}} \end{pmatrix} \Big[D^2_{(\mu,x)}\mathcal{L}(x,\mu^*)\Big] \begin{pmatrix} -\mu^* \\ x - x^* \end{pmatrix}$$

Since $x \in B_\varepsilon(x^*) \cap M$ and $y \in B_\varepsilon(x^*)$, the final term is negative for $x \neq x^*$.

This shows $f(x) < f(x^*)$ for all $x \in B_\varepsilon(x^*)$ with $x \neq x^*$. The point x^* is a strict local maximum on the constraint set M. ∎

19.12 Second Order Sufficient Conditions: Minimum

If the Hessian is positive definite on $T_{\mathbf{x}^*}M = \ker D_{\mathbf{x}^*}\mathbf{h}$ at a critical point $\mathbf{x}^*$, then $\mathbf{x}^*$ is a minimum.

Theorem 19.12.1. *Let $U \subset \mathbb{R}^m$ and suppose $f\colon U \to \mathbb{R}$ and $\mathbf{h}\colon U \to \mathbb{R}^\ell$ are $\mathcal{C}^2$ functions. Suppose further that* rank $D\mathbf{h}(\mathbf{x}^*) = \ell$ *holds (NDCQ) and define the constraint set by $M = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$. The Lagrangian is*

$$\mathcal{L} = f(\mathbf{x}) - \boldsymbol{\mu}^{\mathsf{T}}\big(\mathbf{c} - \mathbf{h}(\mathbf{x})\big).$$

Suppose

1. $\mathbf{x}^* \in M$.
2. *There are $\boldsymbol{\mu}^* \in \mathbb{R}^\ell$ with $D_{\mathbf{x}}\mathcal{L} = \mathbf{0}$ and $D_{\boldsymbol{\mu}}\mathcal{L} = \mathbf{0}$ at $(\mathbf{x}^*, \boldsymbol{\mu}^*)$.*
3. *The Hessian of the Lagrangian with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \boldsymbol{\mu}^*)$, $D^2_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is positive definite on the tangent space $T_{\mathbf{x}^*}M = \ker D_{\mathbf{x}^*}\mathbf{h} = \{\mathbf{v} : \big(D_{\mathbf{x}^*}\mathbf{h}\big)\mathbf{v} = \mathbf{0}\}$.*

Then $\mathbf{x}^*$ is a strict local constrained minimum of f on M.

Moreover, the conclusion remains true if (3) is replaced by (3′).

3′. *The last $(m-\ell)$ leading principal minors of the bordered Hessian $\mathbf{B}$ have the same sign and the determinant of $\mathbf{B}$ obeys $(-1)^\ell \det \mathbf{B} > 0$.*

Proof. Mimic the proof of Theorem 19.10.1. ∎

19.13 2nd Order Conditions: Mixed Constraint Maxima

We'll jump directly to the full mixed constraint version of the theorem.

Theorem 19.13.1. *Let $U \subset \mathbb{R}^m$ and suppose the functions $f\colon U \to \mathbb{R}$, $\mathbf{g}\colon U \to \mathbb{R}^k$, and $\mathbf{h}\colon U \to \mathbb{R}^\ell$ are $\mathcal{C}^2$ functions. Suppose further that there are $\hat{k}$ binding constraints at $\mathbf{x}^*$. Let $\hat{\mathbf{g}}$ be the vector of binding inequality constraints at $\mathbf{x}^*$ with corresponding constants $\hat{\mathbf{b}}$. Set*

$$\mathbf{G}(\mathbf{x}) = \begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}$$

Suppose $\operatorname{rank} D\mathbf{G}(\mathbf{x}^*) = \hat{k} + \ell < m$ *holds (NDCQ) and define the constraint set by $M = \{\mathbf{x} : \mathbf{g}(\mathbf{x}) \le \mathbf{b}, \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$. Form the Lagrangian $\mathcal{L}$:*

$$\mathcal{L} = f(\mathbf{x}) - \boldsymbol{\lambda}^T\big(\hat{\mathbf{g}}(\mathbf{x}) - \hat{\mathbf{b}}\big) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

Suppose

1. $\mathbf{x}^* \in M$.
2. *There are $\boldsymbol{\lambda}^* \in \mathbb{R}^k$ and $\boldsymbol{\mu}^* \in \mathbb{R}^\ell$ with*

$$\begin{gathered} D_{\mathbf{x}}\mathcal{L} = \mathbf{0} \quad \text{at } (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*), \\ \boldsymbol{\lambda}^* \ge \mathbf{0}, \\ \lambda_1^*\big(g_1(\mathbf{x}^*) - b_1\big) = 0, \dots, \lambda_k^*\big(g_k(\mathbf{x}^*) - b_k\big) = 0, \\ \mathbf{h}(\mathbf{x}^*) = \mathbf{c}. \end{gathered}$$

3. *The Hessian of the Lagrangian with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$, $D_{\mathbf{x}}^2\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is negative definite on the tangent space $T_{\mathbf{x}^*}M = \ker D_{\mathbf{x}^*}\mathbf{G} = \{\mathbf{v} : \big(D_{\mathbf{x}^*}\mathbf{G}\big)\mathbf{v} = \mathbf{0}\}$.*

Then $\mathbf{x}^*$ is a strict local constrained maximum of f on M.

Moreover, the conclusion remains true if (3) is replaced by (3′).

3′. *The last $(m - \hat{k} - \ell)$ leading principal minors of the bordered Hessian $\mathbf{B}$ alternate in sign and the determinant of $\mathbf{B}$ obeys $(-1)^m \det \mathbf{B} > 0$.*

19.14 2nd Order Conditions: Mixed Constraint Minima

The version for minima is only slightly different. Although the differences are minor, they make all the difference in the world, resulting in minima rather than maxima.

Theorem 19.14.1. *Let $U \subset \mathbb{R}^m$ and suppose the functions $f\colon U \to \mathbb{R}$, $\mathbf{g}\colon U \to \mathbb{R}^k$, and $\mathbf{h}\colon U \to \mathbb{R}^\ell$ are $\mathcal{C}^2$ functions. Suppose further that there are $\hat{k}$ binding constraints at $\mathbf{x}^*$. Let $\hat{\mathbf{g}}$ be the vector of binding inequality constraints at $\mathbf{x}^*$ with corresponding constants $\hat{\mathbf{b}}$. Set*

$$\mathbf{G}(\mathbf{x}) = \begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}$$

Suppose $\operatorname{rank} D\mathbf{G}(\mathbf{x}^*) = \hat{k} + \ell < m$ *holds (NDCQ) and define the constraint set by* $M = \{\mathbf{x} : \mathbf{g}(\mathbf{x}) \geq \mathbf{b}, \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$. *The Lagrangian is*

$$\mathcal{L} = f(\mathbf{x}) - \boldsymbol{\lambda}^\mathsf{T}\big(\hat{\mathbf{g}}(\mathbf{x}) - \mathbf{b}\big) - \boldsymbol{\mu}^\mathsf{T}\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big).$$

Suppose

1. $\mathbf{x}^* \in M$.
2. *There are $\boldsymbol{\lambda}^* \in \mathbb{R}^k$ and $\boldsymbol{\mu}^* \in \mathbb{R}^\ell$ with*

$$\begin{gathered} D_\mathbf{x}\mathcal{L} = \mathbf{0} \quad \textit{at } (\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*), \\ \boldsymbol{\lambda}^* \geq \mathbf{0}, \\ \lambda_1^*\big(g_1(\mathbf{x}^*) - b_1\big) = 0, \dots, \lambda_k^*\big(g_k(\mathbf{x}^*) - b_k\big) = 0, \\ \mathbf{h}(\mathbf{x}^*) = \mathbf{c}. \end{gathered}$$

3. *The Hessian of the Lagrangian with respect to $\mathbf{x}$ at $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$, $D^2_\mathbf{x}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is positive definite on the tangent space $T_{\mathbf{x}^*}M = \ker D_{\mathbf{x}^*}\mathbf{G} = \{\mathbf{v} : \big(D_{\mathbf{x}^*}\mathbf{G}\big)\mathbf{v} = \mathbf{0}\}$.*

Then $\mathbf{x}^*$ is a strict local constrained minimum of f on M.

Moreover, the conclusion remains true if (3) is replaced by (3′).

3′. *The last $(m - \hat{k} - \ell)$ leading principal minors of the bordered Hessian $\mathbf{B}$ have the same sign and the determinant of $\mathbf{B}$ obeys $(-1)^{\hat{k}+\ell}\det\mathbf{B} > 0$.*

19.15 Second Order Necessary Conditions, Preliminaries

We will limit ourselves to a single theorem giving necessary conditions for a local maximum. The local minimum case is similar. This involves semidefinite matrices. There are also sufficient conditions using semidefinite matrices, analogous to Theorem 17.4.1 and Theorem 17.4.1 Restated.

Let $\hat{k}$ be the number of binding inequality constraints at $\mathbf{x}^*$ and $\hat{\mathbf{g}}$ the corresponding vector of binding inequality constraints. Form the $\hat{k}+\ell$-vector of binding constraints at $\mathbf{x}^*$,

$$\mathbf{G}(\mathbf{x}) = \begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix}$$

Consider the bordered Hessian of the Lagrangian:

$$\mathbf{B} = \begin{pmatrix} 0_{(\hat{k}+\ell)\times(\hat{k}+\ell)} & D\mathbf{G}(\mathbf{x}^*) \\ D\mathbf{G}(\mathbf{x}^*)^T & D^2_{\mathbf{x}^*}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \end{pmatrix}$$

Bordered Principal Submatrices. Let $\hat{\mathbf{B}}_r$ denote any $r \times r$ principal submatrix with $r > \hat{k}+\ell$, containing the $0_{(\hat{k}+\ell)\times(\hat{k}+\ell)}$ block in upper left hand corner of $\mathbf{B}$. We will refer to these as *bordered principal submatrices*.

19.16 Second Order Necessary Conditions

Theorem 19.16.1. *Let $U \subset \mathbb{R}^m$ and suppose the functions $f\colon U \to \mathbb{R}$, $\mathbf{g}\colon U \to \mathbb{R}^k$, and $\mathbf{h}\colon U \to \mathbb{R}^\ell$ are $\mathcal{C}^2$ functions. Suppose further that there are $\hat{k}$ binding inequality constraints at $\mathbf{x}^*$. Let $\hat{\mathbf{g}}$ be the vector of binding inequality constraint functions at $\mathbf{x}^*$ with corresponding constants $\hat{\mathbf{b}}$. Then set*

$$\mathbf{G}(\mathbf{x}) = \begin{pmatrix} \hat{\mathbf{g}}(\mathbf{x}) \\ \mathbf{h}(\mathbf{x}) \end{pmatrix},$$

which includes all binding constraints at $\mathbf{x}^*$.

Suppose that $\operatorname{rank} D\mathbf{G}(\mathbf{x}^*) = \hat{k} + \ell$ holds (NDCQ) and define the constraint set by $M = \{\mathbf{x} : \mathbf{g}(\mathbf{x}) \le \mathbf{b}, \mathbf{h}(\mathbf{x}) = \mathbf{c}\}$. Suppose that $\mathbf{x}^* \in U^0$ is a local maximum of f on M, that $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is a critical point for the Lagrangian

$$\mathcal{L} = f(\mathbf{x}) - \boldsymbol{\lambda}^T\big(\hat{\mathbf{g}}(\mathbf{x}) - \hat{\mathbf{b}}\big) - \boldsymbol{\mu}^T\big(\mathbf{h}(\mathbf{x}) - \mathbf{c}\big),$$

and that the complementary slackness and non-negativity conditions hold.

Then the Hessian of the Lagrangian at $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is negative semidefinite on the tangent space $T_{\mathbf{x}^*} = \ker D\mathbf{G}(\mathbf{x}^*) = \{\mathbf{v} \in \mathbb{R}^{h+\ell} : \big(D\mathbf{G}(\mathbf{x}^*)\big)\mathbf{v} = \mathbf{0}\}$.

Equivalently, either the minors $\det \hat{\mathbf{B}}_r$ of any given size $r > m-(\hat{k}+\ell)$ are either all non-positive, or all non-negative. The generalized sign alternates with r, and $(-1)^{\hat{k}+\ell} \det \hat{\mathbf{B}} > 0$.

Minimum. If $\mathbf{x}^*$ is a minimum, we have to reverse the inequalities for $\mathbf{g}$, and adjust the Lagrangian accordingly. In that case, we can conclude that the Hessian of the Lagrangian at $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is positive semidefinite on the tangent space $T_{\mathbf{x}^*}$.

Equivalently, the minors $\det \hat{\mathbf{B}}_r$ with $r > m - (\hat{k} + \ell)$ are either all non-positive, or all non-negative. and $(-1)^{m-\hat{k}-\ell} \det \hat{\mathbf{B}} > 0$.

Simon and Blume punted on the determinant conditions. The best thing to do is to look in the source. Our determinant condition is a rewritten version of that in Debreu (1952).[1]

[1] Debreu, Gerard (1952) Definite and semidefinite quadratic forms, *Econometrica*, **20**, 295–300.

19.17 Parameters and Optima

Earlier, when considering the Envelope Theorem and related results, we required that the maximum or minimum be a $\mathcal{C}^1$ function of various parameters. One weakness of that approach is that we gave no method of ensuring that the functions were $\mathcal{C}^1$.

We are ready to remedy that and will give conditions on the primitives of the model, the objectives and constraints, that guarantee the solutions are $\mathcal{C}^1$. The Implicit Function Theorem is the key tool for this.

We start with an unconstrained problem to explore the basic ideas. Consider a simple model with a single parameter a, and two variables, $(x, y) \in U \subset \mathbb{R}^2$. We consider the maximization problem

$$F(a) = \max_{(x,y)\in U} f(x, y, a).$$

If U is compact, the Weierstrass Theorem guarantees a solution. If we have an interior maximum, we know that the first order conditions must be satisfied,

$$\frac{\partial f}{\partial x} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y} = 0.$$

These equations may implicitly define functions $x^*(a)$ and $y^*(a)$ that maximize f for at any given a.

For this to work, the Implicit Function Theorem requires that the (x, y)-derivative of this vector equation be non-singular at $\big(x^*(a), y^*(a), a\big)$. That is, the Hessian matrix

$$D^2_{(x,y)} f\big(\mathbf{x}(a), y(a), a\big)$$

must be non-singular.

If the second order sufficient conditions for a strict local maximum are satisfied, the Hessian $D^2_{(x,y)} f$ will be negative definite, and hence non-singular.

So let's write a theorem summing up what we proved. The Implicit Function Theorem requires that our Implicit Function be defined by a function that is itself $\mathcal{C}^1$. It also allows multiple exogenous variables, like our variable a.

Differentiability is a local property, so the fact that the Implicit Function Theorem only gives local results should not be a problem.

If we consider a similar minimization problem, the only real difference is that the similar second order sufficient condition is that $D^2_{\mathbf{x}} f$ be positive definite at $\mathbf{x}(a)$.

Let's add these considerations to our theorem.

19.18 Smoothness of Unconstrained Optima

Here's a version of such a theorem, stated for $\mathbf{x} \in \mathbb{R}^m$.

Theorem 19.18.1. *Let* $U \subset \mathbb{R}^m$ *and* $V \subset \mathbb{R}$. *Suppose* $f\colon U \times V \to \mathbb{R}$ *is* $\mathcal{C}^2$ *in* $\mathbf{x} \in U$ *and* $\mathcal{C}^1$ *in* $\mathbf{a} \in V$. *Suppose also there are* $(\mathbf{x}_0, a_0) \in U^0 \times V^0$ *with* $D_{\mathbf{x}} f(\mathbf{x}_0, \mathbf{a}_0) = 0$.

If the Hessian $D_{\mathbf{x}}^2 f(\mathbf{x}_0, a_0)$ *is either positive definite or negative definite, then there is an* $\varepsilon > 0$ *and a* $\mathcal{C}^1$ *function* $\mathbf{x}^*\colon B_\varepsilon(a_0) \to \mathbb{R}^m$ *with* $\mathbf{x}^*(a)$ *obeying*

$$D_{\mathbf{x}} f\big(\mathbf{x}(\mathbf{a}), \mathbf{a}\big) = 0$$

on $B_\varepsilon(\mathbf{a}_0)$ *with* $\mathbf{x}(\mathbf{a}_0) = \mathbf{x}_0$.

The function maximizes f *if the Hessian is negative definite and minimizes* f *if the Hessian is positive definite.*

Proof. Left as an exercise to the reader. All of the key components were mentioned on the previous page on included in previous sufficient conditions.

19.19 Smoothness of Constrained Optima

It's not much harder to show the solution to optimization problems with equality constraints are smooth.

Consider the problem with ℓ equality constraints.

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}, a) \\ &\text{s.t } \mathbf{h}(\mathbf{x}, a) = 0 \end{aligned} \tag{19.19.12}$$

where $f\colon U \to \mathbb{R}$ and $\mathbf{h}\colon V \to \mathbb{R}^\ell$ with $U, V \subset \mathbb{R}^m$.

If the NDCQ holds, $\operatorname{rank} D_{\mathbf{x}}\mathbf{h} = \ell$ at the optimal $\mathbf{x}^*(a)$.

The Lagrangian is

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, a) = f(\mathbf{x}, a) - \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}, a)$$

and any constrained optimum, $\mathbf{x}^*(a)$ satisfies the first order conditions

$$D_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, a) = \mathbf{0}, \quad D_{\boldsymbol{\mu}}\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, a) = \mathbf{0}.$$

This is a system of $m + \ell$ equations (m from the first group, ℓ from the second). There are also $m + \ell$ endogenous variables $\mathbf{x}$ and $\boldsymbol{\mu}$, along with one exogenous variable a. Provided we know the problem has a solution (e.g., by Weierstrass's Theorem) the Implicit Function Theorem tells us the solution is $\mathcal{C}^1$ if the derivative of this system with respect to the endogenous variables is non-singular.

We compute

$$D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L} = \begin{pmatrix} \mathbf{0}_{\ell\times\ell} & -D\mathbf{h} \\ & \\ -D\mathbf{h}^T & D^2_{\mathbf{x}}\mathcal{L} \end{pmatrix}. \tag{19.19.13}$$

It is the version of the bordered Hessian we previously encountered following Theorem 19.10.1. This matrix will be singular if NDCQ fails. That means that if we require $D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}$ be invertible, we are also requiring that the NDCQ is satisfied, as shown by the following theorem.

Theorem 19.19.1. *Suppose $D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}(\mathbf{x}^*)$ given by equation (19.19.13) is invertible. Then* $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$.

Proof. Consider

$$\left[D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}(\mathbf{x}^*)\right]\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} = \begin{pmatrix} -\big(D\mathbf{h}(\mathbf{x}^*)\big)\mathbf{v} \\ -\big(D\mathbf{h}(\mathbf{x}^*)\big)^T\mathbf{u} + \left[D^2_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*)\right]\mathbf{v} \end{pmatrix}$$

for any $\mathbf{u} \in \mathbb{R}^k$ and $\mathbf{v} \in \mathbb{R}^m$. If $D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}$ is invertible, $\ell + m \le \operatorname{rank} D\mathbf{h}(\mathbf{x}^*) + m \le \ell + m$, so we must have $\operatorname{rank} D\mathbf{h}(\mathbf{x}^*) = \ell$. ∎

19.20 Smooth Parameterization Theorem

The above reasoning leads to the following theorem.

Theorem 19.20.1. *Let $\mathbf{x}^*(a)$ be the solution to the parameterized constrained maximization problem in equation (19.19.12) and let $\boldsymbol{\mu}^*(a)$ be the corresponding vector of Lagrange multipliers. Fix a_0. If the bordered Hessian matrix $D^2_{(\boldsymbol{\mu},\mathbf{x})}\mathcal{L}$ is non-singular at $\big(\mathbf{x}^*(a_0), \boldsymbol{\mu}^*(a_0), a_0\big)$, then*

(a) *the NDCQ holds at $\big(\mathbf{x}^*(a_0), \boldsymbol{\mu}^*(a_0), a_0\big)$, and*

(b) *$\mathbf{x}^*(a)$ and $\boldsymbol{\mu}^*(a)$ are $\mathcal{C}^1$ at $a = a_0$.*

This theorem can also be generalized to a vector of parameters, and to cover cases with inequality constraints. For the latter, its necessary to focus on the binding constraints. Rather than considering a collection of variations on a theorem, we return to the issue of constraint qualification.

19.21 Alternatives to NDCQ Constraint Qualification

There are a number of alternate ways to handle the constraint qualification problem. Some of them will work in cases where the standard method does not, as in the example in section 18.26.

One alternative is to use an additional multiplier for the objective.

Theorem 19.21.1. *Let f and h be $\mathcal{C}^1$ functions on $U \subset \mathbb{R}^2$. Suppose that $\mathbf{x}^*$ solves*

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}) \\ &\quad s.t\ h(\mathbf{x}) = c \end{aligned}$$

Form the Lagrangian

$$\mathcal{L}(\mathbf{x}, \mu_0, \mu_1) = \mu_0 f(\mathbf{x}) - \mu_1\big(h(\mathbf{x}) - c\big),$$

by including a multiplier for the objective function f. Then there exist multipliers μ_0^* and μ_1^* such that

(a) *μ_0^* and μ_1^* are not both zero,*
(b) *μ_0^* is either 0 or 1,*
(c) *$(\mathbf{x}^*, \boldsymbol{\mu}^*)$ satisfies the first order conditions*

$$\begin{aligned} \mathbf{0} &= D_{\mathbf{x}}\mathcal{L} = \mu_0 D_{\mathbf{x}} f - \mu_1 D_{\mathbf{x}} h \\ 0 &= \frac{\partial \mathcal{L}}{\partial \mu_1} = c - h(\mathbf{x}). \end{aligned}$$

Proof. If $D_{\mathbf{x}} h \neq \mathbf{0}$, constraint qualification (NDCQ) is satisfied, and there is a solution with $\mu_0^* = 1$ by Theorem 18.7.1.

If $D_{\mathbf{x}} h = \mathbf{0}$, then set $\mu_0^* = 0$. ∎

In this case, the extra multiplier μ_0 serves as a device to remind us to check the critical points of the constraints as well as the objective.

19.22 Example 18.26 Revisited

Recall that we considered the problem

$$\begin{aligned} &\max_{(x,y)} x \\ &\text{s.t. } x^3 + y^2 = 0. \end{aligned}$$

Let's try Theorem 19.21.1 on it. Set up the Lagrangian

$$\mathcal{L} = \mu_0 x - \mu_1(x^3 + y^2).$$

The first order conditions become

$$\begin{aligned} \mu_0 &= 3\mu_1 x^2 \\ 0 &= -2\mu_1 y \\ 0 &= x^3 + y^2. \end{aligned}$$

We saw in section 18.26 that when $\mu_0 = 1$, there is no solution. What about $\mu_0 = 0$. Then $0 = 3\mu_1 x^2$, so either $x = 0$ or $\mu_1 = 0$.

If $x < 0$, $\mu_1 = 0$, which is forbidden by the condition that at least one μ_i must be non-zero.

Finally, if $x = 0$, $y = 0$, and any value of μ_1 will satisfy the first order equations. This yields the correct solution $(x, y) = (0, 0)$.

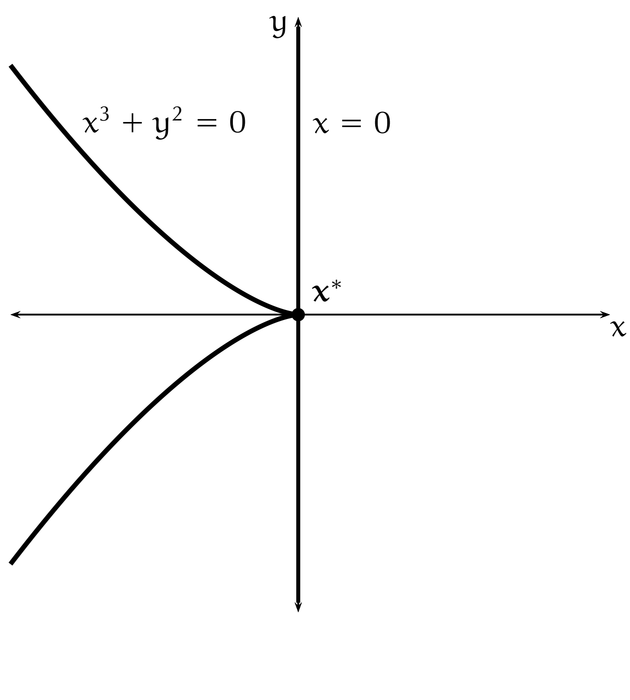

Figure 19.22.1: The solution is at the cusp, $\mathbf{x}^* = (0, 0)$. There $Dh(\mathbf{x}^*) = (0, 0)$, forcing the multiplier $\mu_0 = 0$ because $Df = (1, 0) \neq (0, 0)$.

19.23 Fritz John Theorem

This version of Theorem 19.21.1, including a multiplier on the objective, allows for inequality constraints.

Theorem 19.23.1. *Let $U \subset \mathbb{R}^m$. Suppose $f\colon U \to \mathbb{R}$ and $\mathbf{g}\colon U \to \mathbb{R}^k$ are $\mathcal{C}^1$ functions. Suppose further that $\mathbf{x}^*$ is a local max of f under the constraints*

$$\mathbf{g}(\mathbf{x}) \le \mathbf{b}.$$

Form the Lagrangian

$$\mathcal{L}(\mathbf{x}, \lambda_0, \dots, \lambda_k) = \lambda_0 f(\mathbf{x}) - \sum_{j=1}^{k} \lambda_j \big(g_j(\mathbf{x}) - b_j\big)$$

with a multiplier λ_0 for the objective function. Then there exist $\lambda_0^*, \dots, \lambda_k^*$ such that:

1. $D_x \mathcal{L}(\mathbf{x}^*, \lambda_0^*, \dots, \lambda_k^*) = \mathbf{0}$.
2. $\lambda_1^*[g_1(\mathbf{x}^*) - b_1], \dots, \lambda_k^*[g_k(\mathbf{x}^*) - b_k] = 0$.
3. $\lambda_1^* \ge 0, \dots, \lambda_k^* \ge 0$.
4. $\mathbf{g}(\mathbf{x}^*) \le \mathbf{b}$.
5. $\lambda_0^* = 0$ *or* $\lambda_0^* = 1$, *and*
6. $(\lambda_0^*, \dots, \lambda_k^*) \ne \mathbf{0}$.

19.24 More Constraint Qualification: Maximia

There are other constraint qualification conditions that can be used. The following theorem collects some of the them without proof.

Theorem 19.24.1. *Let $U \subset \mathbb{R}^m$ and suppose $f\colon U \to \mathbb{R}$ and $\mathbf{g}\colon U \to \mathbb{R}^k$ are $\mathcal{C}^1$ functions. Suppose also that $\mathbf{x}^*$ is a local maximum of f under the constraints*

$$\mathbf{g}(\mathbf{x}) \le \mathbf{b}.$$

Let $\hat{\mathbf{g}}(\mathbf{x}) \in \mathbb{R}^h$ denote the vector of binding constraints at $\mathbf{x}^*$, and suppose $\hat{\mathbf{g}}$ obeys one of the following:

(a) NCCQ: *The derivative $D\hat{\mathbf{g}}(\mathbf{x}^*)$ has rank h.*
(b) Karush-Kuhn-Tucker CQ: *For any $\mathbf{v} \in \mathbb{R}^m$ obeying $D\hat{\mathbf{g}}(\mathbf{x}^*)\mathbf{v} \le \mathbf{0}$ there is an $\varepsilon > 0$ and a $\mathcal{C}^1$ curve $\alpha\colon [0,\varepsilon) \to \mathbb{R}^m$ such that:*
   (i) $\alpha(0) = \mathbf{x}^*$,
   (ii) $\alpha'(0) = \mathbf{v}$, *and*
   (iii) $g_i\big(\alpha(t)\big) \le b_i$ *for* $i = 1, \dots, k$ *and all* $t \in [0, \varepsilon)$.
(c) Slater CQ: *There is a ball $V \subset \mathbb{R}^m$ containing $\mathbf{x}^*$ with the g_i convex on V and there exists a point $\mathbf{z} \in V$ with $\mathbf{g}(\mathbf{z}) \ll \mathbf{b}$.*
(d) Concave CQ: *The g_i are concave functions.*
(e) Linear CQ: *The g_i are linear functions.*

Then we can set $\lambda_0^* = 1$ in Theorem 19.23.1.

The first constraint qualification condition is our old friend, the NDCQ. The second was used by Kuhn and Tucker in their original work on optimization with inequality constraints. The Slater condition applies with convex constraints and requires that the constraint set have a non-empty interior. The concave CQ ensures the constraint set is convex. The last condition derives from linear programming models. It can be considered a special case of the Concave CQ.

19.25 More Constraint Qualification: Minima

Some of the constraint qualification conditions work a bit differently for minima. Notice that the way the constraints are written has been reversed in the following theorem. Further, the KKTCQ, Slater CQ, and Concave (now Convex) CQ are all modified.

Theorem 19.25.1. *Let* $U \subset \mathbb{R}^m$ *and suppose* $f\colon U \to \mathbb{R}$ *and* $\mathbf{g}\colon U \to \mathbb{R}^k$ *are* $\mathcal{C}^1$ *functions. Suppose also that* $\mathbf{x}^*$ *is a local minimum of* f *under the constraints*

$$\mathbf{g}(\mathbf{x}) \geq \mathbf{b}.$$

Let $\hat{\mathbf{g}}(\mathbf{x}) \in \mathbb{R}^h$ *denote the vector of binding constraints at* $\mathbf{x}^*$*, and suppose* $\hat{\mathbf{g}}$ *obeys one of the following:*

(a) NCCQ: *The derivative* $D\hat{\mathbf{g}}(\mathbf{x}^*)$ *has rank* h*.*

(b) Karush-Kuhn-Tucker CQ: *For any* $\boldsymbol{v} \in \mathbb{R}^m$ *obeying* $D\hat{\mathbf{g}}(\mathbf{x}^*)\boldsymbol{v} \geq \mathbf{0}$ *there is an* $\varepsilon > 0$ *and a* $\mathcal{C}^1$ *curve* $\alpha\colon [0,\varepsilon) \to \mathbb{R}^m$ *such that:*

   (i) $\alpha(0) = \mathbf{x}^*$*,*

   (ii) $\alpha'(0) = \boldsymbol{v}$*, and*

   (iii) $g_i\big(\alpha(t)\big) \geq b_i$ *for* $i = 1, \dots, k$ *and all* $t \in [0, \varepsilon)$*.*

(c) Slater CQ: *There is a ball* $V \subset \mathbb{R}^m$ *containing* $\mathbf{x}^*$ *with each* g_i *is concave on* V *and there exists a point* $\boldsymbol{z} \in V$ *with* $\mathbf{g}(\boldsymbol{z}) \gg \mathbf{b}$*.*

(d) Convex CQ: *The* g_i *are convex functions.*

(e) Linear CQ: *The* g_i *are linear functions.*

Then we can set $\lambda_0^* = 1$ *in Theorem 19.23.1.*

November 6, 2020