4. Nonlinear Programming Problems and the Kuhn Tucker conditions (cont) 1

4.4. KT conditions and the Lagrangian approach 1

4.5. Interpretation of the Lagrange Multiplier 3

4.6. A worked solution to an NPP: S&B #18.18 (on the problem set) 6

4.7. Computing a solution to a NPP: a simple worked example 10

4.8. Computed solution to a NPP: ARE problem set example. 13

4.9. Second Order conditions Without Quasi-Concavity 21

4.9.1. One linear equality constraint 22

4.9.2. One nonlinear equality constraint 25

4. Nonlinear Programming Problems and the Kuhn Tucker conditions (cont)

4.4. KT conditions and the Lagrangian approach

There’s an alternative way of writing the KT conditions, which may be more familiar. Set up the Lagrangian function, take its FOC’s, and look for a solution to them. The FOC for the Lagrangian will be identical to the KT conditions. Why bother with the Lagrangian? Absolutely no reason why you shouldn’t just look at the KT conditions, and write down the conditions explicitly for a given problem. But most people find that approach a bit abstract. So we offer this less abstract route. Important to see that you end up at exactly the same place, though.
The Lagrangian function for the general NPP is as follows:

\[ L(x, \lambda) = f(x) + \lambda(b - g(x)) \]

\[ = f(x) + \sum_{j=1}^{m} \lambda_j (b_j - g^j(x)) \]

Look for an \( \bar{x} \) and \( \bar{\lambda} \geq 0 \) satisfying the following conditions

\[ \partial L(\bar{x}, \bar{\lambda}) / \partial x_i = 0; \quad \partial L(\bar{x}, \bar{\lambda}) / \partial \lambda_j \geq 0; \quad \bar{\lambda}_j \partial L(\bar{x}, \bar{\lambda}) / \partial \lambda_j = 0. \]

Note well: this specification of the Lagrangian first order conditions differs from many textbooks, eg., Chiang. this books impose the condition \( \partial L(\bar{x}, \bar{\lambda}) / \partial x_i \leq 0 \). The source of the difference is that they impose an additional restriction on their specification of the NPP, i.e., that \( x \geq 0 \). My specification is more general: if you want \( x \geq 0 \), include this as one of your constraints.

To see that these are the same as the K-T conditions, do the derivatives one by one

\[ \partial f(\bar{x}) / \partial x_i = \sum_{j=1}^{m} \lambda_j \cdot \partial g^j(\bar{x}) / \partial x_i \]

The inner product you are calculating is precisely the inner product you calculate when you pre-multiply the \( \lambda^T \) vector with the \( i \)'th column of the Jacobian matrix.

The following is very important for what follows: the maximized value of the objective function is identically equal to the value of the Lagrangian at the solution to the NPP. More formally:

**Theorem:** At a solution \( \bar{x} \) to the NPP, \( L(\bar{x}, \bar{\lambda}) = f(\bar{x}) \).
To see that this is correct note that \( L(x, \lambda) = f(x) + \lambda(b - g(x)) \). Each component of the inner product \( \lambda(b - g(x)) \) is zero, since \( \text{either } (b_j - g^i(x)) \text{ is zero or } \lambda_j \text{ is zero} \). Hence the entire second term on the right hand side of the Lagrangian is zero.

4.5. **Interpretation of the Lagrange Multiplier**

The Lagrangian \( \lambda_j \) has an interpretation that proves to be important in a wide variety of economic applications.

- it is a measure of how much bang for the buck you get when you relax the \( j \)’th constraint. That is, if you increase \( b_j \) by one unit, then the maximized value of your objective function will go up by \( \lambda_j \) units.
- Example: think of maximizing utility on a budget constraint. the higher are prices the longer is the gradient vector, and so the shorter is \( \lambda \), i.e., the less bang for the buck you get, literally. I.e., relaxing your budget constraint by a dollar doesn’t by much more of a bundle, so that your utility can’t go up by much. On the other hand, for a fixed price vector, the length of the gradient of your utility function is a measure of how easy to please you are. If you are a kid who gets a lot of pleasure out of penny candy, then relaxing the budget constraint by a dollar will buy you a lot.
- the importance of this in economic applications is that you often want to know what the economic cost of a constraint is, e.g., suppose you are maximizing output subject to a resource constraint: what’s it worth to you to increase the level of available resources by one unit. Get answer by looking at the Lagrangian.
The mathematical proof is as usual completely trivial. Let \( M(b) \) denote the maximized value of the objective function when the constraint vector is \( b \).

\[
M(b) = f(\bar{x}(b)) = L(\bar{x}(b), \bar{\lambda}(b)) = f(x) + \lambda(b - g(x)).
\]

we have

\[
\frac{dM(b)}{db_j} = \frac{dL(\bar{x}(b), \bar{\lambda}(b))}{db_j} = \sum_{i=1}^{n} \left\{ f_i(\bar{x}(b)) - \sum_{k=1}^{m} \lambda_k g_k^i(\bar{x}(b)) \right\} \frac{dx_i}{db_j} + \sum_{k=1}^{m} \frac{d\lambda_k}{db_j} (b_k - g^k(x)) + \lambda_j
\]

– For each \( i \), the term in curly brackets is zero, by the KT conditions.

– For each \( k \),

  * If \((b_k - g^k(x)) = 0\), then \( \frac{d\lambda_k}{db_k} (b_k - g^k(x)) \) is zero.

  * If \((b_k - g^k(x)) < 0\), then \( \lambda_k(\cdot) \) will be zero on a neighborhood of \( b \), so that \( \frac{d\lambda_k(\cdot)}{db_k} = 0 \).

– The only term remaining is \( \lambda_j \).

– Conclude that \( \frac{dM(b)}{db_j} = \lambda_j \)

Here’s a more intuitive approach for the case of a single binding constraint.

– increase the constraint \( b \) to \( b + db \).

– you should respond by moving from solution \( x \) in the direction of steepest ascent, i.e., in the direction that the gradient is pointing.

– move from \( x \) in the direction that \( \nabla f(x) \) is pointing until you reach the new constraint, i.e., define \( dx \) by \( g(x + dx) = b + db \). Now initially we had \( g(x) = b \). Hence

\[
db = g(x + dx) - g(x) \approx \nabla g(x) dx
\]
while

\[ df = \nabla f(x)\, dx \]

which by the KT conditions

\[ = \lambda \nabla g(x)\, dx \]

Hence

\[ df = \lambda db \]

Hence (being a little fast and loose with infinitesimals), \( df/db = \lambda \).

Note that \( \lambda \) will be larger

- the more rapidly \( f \) increases with \( x \) (i.e., the longer is the vector \( \nabla f(x) \)).
- the less rapidly \( g \) increases with \( x \) (i.e., the shorter is the vector \( \nabla g(x) \)).

Note also that the above “proves” the 2nd part of the mantra, i.e., if constraint \( j \) is satisfied with equality but is not binding, then the weight on this constraint must be zero when you write \( \nabla f(x) \) as a nonnegative linear combination of the \( \nabla g^j(x) \)'s. The following argument proves this rather sloppily: (a) if constraint \( j \) is satisfied with equality but is not binding, then by definition \( \frac{df}{db_j} = 0 \); (b) since \( \frac{df}{db_j} = \lambda_j \), then \( \lambda_j \) must be zero.
4.6. A worked solution to an NPP: S&B #18.18 (on the problem set)

The example: S&B qu 18.18

\[
\begin{align*}
\min 2x^2 + 2y^2 - 2xy - 9y & \quad \text{s.t.} \\
4x + 3y & \leq 10 \\
y - 4x^2 & \geq -2 \\
x & \geq 0 \\
y & \geq 0
\end{align*}
\]

Flip the signs so it’s a max problem, respecify nonnegativity constraints as inequalities

\[
\begin{align*}
\max f(x, y) &= 2xy + 9y - 2x^2 - 2y^2 \quad \text{s.t.} \\
4x + 3y & \leq 10 \\
4x^2 - y & \leq 2 \\
-x & \leq 0 \\
-y & \leq 0
\end{align*}
\]

Now set up Lagrangian:

\[
L(x, y, \lambda) = (2xy + 9y - 2x^2 - 2y^2) + \lambda_1(10 - 4x - 3y) + \lambda_2(2 - 4x^2 + y) + \lambda_3x + \lambda_4y
\]
Recall from (1), that the first order conditions were

\[ \frac{\partial L(\bar{x}, \bar{\lambda})}{\partial x_i} = 0; \quad \frac{\partial L(\bar{x}, \bar{\lambda})}{\partial \lambda_j} \geq 0; \quad \bar{\lambda}_j \frac{\partial L(\bar{x}, \bar{\lambda})}{\partial \lambda_j} = 0. \]  

In this particular problem, these conditions imply

\[ L_x = 2y - 4x - 4\lambda_1 - 8x\lambda_2 + \lambda_3 = 0 \]
\[ L_y = 2x - 4y + 9 - 3\lambda_1 + \lambda_2 + \lambda_4 = 0 \]
\[ L_{\lambda_1} = 10 - 4x - 3y \geq 0 \]
\[ L_{\lambda_2} = 2 - 4x^2 + y \geq 0 \]
\[ L_{\lambda_3} = x \geq 0 \]
\[ L_{\lambda_4} = y \geq 0 \]
3) Assume at least one nonnegativity constraint satisfied with equality:

(a) First assume we’re at the origin. \( x = y = 0; \)
- \( L\lambda_1 > 0; L\lambda_2 > 0 \)
- \( \lambda_1 = \lambda_2 = 0 \)
- \( L_x = 0 - 0 - 0 - 0 + \lambda_3 = 0 \Rightarrow \lambda_3 = 0; \)
- \( L_y = 0 - 0 + 9 - 0 + 0 + \lambda_4 > 0. \) \( \otimes \)

(b) Now assume just \( x \) is positive \( x > y = 0; \)
- \( \lambda_3 = 0; \)
- \( L\lambda_2 = 2 - 4x^2 \geq 0 \Rightarrow x \leq \sqrt{0.5} \)
- \( L\lambda_3 > 10 - 4\sqrt{0.5} > 0 \Rightarrow \lambda_1 = 0 \)
- \( L_y = 2x + 9 + \lambda_2 + \lambda_4 > 0. \) \( \otimes \)

(c) Now assume just \( y \) is positive \( y > x = 0; \)
- \( \lambda_4 = 0; \)
- \( L\lambda_2 = 2 + y > 0 \)
- \( \lambda_2 = 0; \)
- \( L_x = 2y + \lambda_3 - 4\lambda_1 = 0 \Rightarrow \lambda_1 > 0 \Rightarrow y = 10/3; \)
- \( L_y = -40/3 + 9 - 3\lambda_1 < 0; \) \( \otimes \)

4) Conclude that \( x > 0; y > 0; \lambda_3 = \lambda_4 = 0. \)

5) Assume both \( x \) and \( y \) are positive:

(a) \( L\lambda_1 = L\lambda_2 = 0 \)
- \( \lambda_3 = \lambda_4 = 0 \) (because both \( x \) and \( y \) are positive).
- \( L\lambda_1 = 10 - 4x - 3y = 0 \Rightarrow y = (10 - 4x)/3 \)
- \( L\lambda_2 = 2 - 4x^2 + (10 - 4x)/3 = 0 \Rightarrow 3x^2 + x - 4 = 0. \)
- i.e., \( (3x + 4) \cdot (x - 1) = 0 \Rightarrow x = 1 \Rightarrow y = 2. \)
- \( L_y = 4 - 8 + 9 > 0; \) \( \otimes \)
(b) $L_{\lambda_1} > 0; L_{\lambda_2} = 0$ (only the quadratic constraint satisfied with equality)

- $\lambda_1 = \lambda_3 = \lambda_4 = 0$
- $L_{\lambda_2} = 2 - 4x^2 + y = 0 \implies y = 4x^2 - 2$;
- $L_x = 0 \implies y \geq 2x$ (otherwise $L_x < 0$).
- $y = 4x^2 - 2$ and $y \geq 2x \implies x \geq 1 \implies y \geq 2$.
- $x \geq 1, y \geq 2 \implies L_{\lambda_2} = 10 - 4x - 3y \leq 0$; ⊗

(c) $L_{\lambda_1} > 0; L_{\lambda_2} > 0$ (solution in the interior of constraint set)

- $\lambda = 0$; (i.e. the whole vector zero)
- $L_x = 0 \implies y = 2x$;
- $L_y = 0 \implies 6x = 9 \implies x = 1.5$;
- $y = 3$;
- $L_{\lambda_1} = 10 - 4 \ast 1.5 - 3 \ast 3 = -5$; ⊗

6) Computing the solution: $L_{\lambda_1} = 0; L_{\lambda_2} > 0$ (only the linear constraint satisfied with equality)

- $\lambda_2 = \lambda_3 = \lambda_4 = 0$
- $L_{\lambda_1} = 10 - 4x - 3y = 0 \implies y = (10 - 4x)/3$
- $L_x = 2(10 - 4x)/3 - 4x - 4\lambda_1 = 0$ or $20 - 20x - 12\lambda_1 = 0$
- $L_y = 2x - 4(10 - 4x)/3 + 9 - 3\lambda_1 = 0$ or $22x - 13 - 9\lambda_1 = 0$

Two equations in two unknowns, i.e.,

\[
\begin{bmatrix}
20 & 12 \\
22 & -9
\end{bmatrix}
\begin{bmatrix}
x \\
\lambda_1
\end{bmatrix}
= 
\begin{bmatrix}
20 \\
13
\end{bmatrix}
\]

so $x = 0.7568; \lambda_1 = 0.4054$;

Check that we have the answer using the mantra:

\[
\nabla f = [2y - 4x \ 2x - 4y + 9]
\]

\[
= [1.6216 \ 1.2162] \propto [4 \ 3] \nabla g = [4 \ 3]
\]
4.7. Computing a solution to a NPP: a simple worked example

How do you actually solve an NPP? Answer is: a process of elimination. You check all the possibilities to see if you can find a point that satisfies the KT conditions, and then you eliminate anything that fails this test. Here you are using the fact that the KT conditions are necessary for a solution, i.e., if they fail this test, they can’t be a maximum. Once you’ve found something that does satisfy the KT conditions, then you have to go back and check that the second order conditions are satisfied.

The example:

$$\max f(x) = (x_1 + 2)(x_2 - 2) \text{ s.t.}$$

$$p_1 x_1 + p_2 x_2 \leq y;$$

$$x_i \geq 0;$$

In this case, $g$ is the matrix above, i.e.,

$$\begin{bmatrix} p_1 & p_2 \\ -1 & 0 \\ 0 & -1 \end{bmatrix};$$

Check the nonvanishing gradient condition: $\nabla f(x) = 0$ iff $x_1 = -2, x_2 = 2$. Clearly this point is outside the constraint set, so we know the gradient is nonvanishing on the constraint set.

Set up Lagrangian:
\( L(x, \lambda) = (x_1 + 2)(x_2 - 2) + \lambda_0(y - p_1 x_1 - p_2 x_2) + \lambda_1 x_1 + \lambda_2 x_2 \)

Recall from (1), that the first order conditions were

\[
\partial L(\bar{x}, \bar{\lambda}) / \partial x_i = 0; \quad \partial L(\bar{x}, \bar{\lambda}) / \partial \lambda_j \geq 0; \quad \bar{\lambda}_j \partial L(\bar{x}, \bar{\lambda}) / \partial \lambda_j = 0.
\]

In this particular problem, these conditions imply

\[
\begin{align*}
L_{x_1} &= x_2 - 2 - \lambda_0 p_1 + \lambda_1 = 0 \\
L_{x_2} &= x_1 + 2 - \lambda_0 p_2 + \lambda_2 = 0 \\
L_{\lambda_0} &= y - p_1 x_1 - p_2 x_2 \geq 0 \\
L_{\lambda_1} &= x_1 \geq 0 \\
L_{\lambda_2} &= x_2 \geq 0.
\end{align*}
\]

Observe that the last three FOC give you back precisely the constraint conditions.

We will set \( p_1 = p_2 = y \) and solve explicitly for a solution. Under this condition, the solution will be at a corner. *NOTE WELL: this solution depends on the particular specification of parameters.* *In general, you could get a solution on the face of the budget line.*

Go through the interior, faces and vertices of the constraint set in turn. (Emphasize that while I can tell by inspection the solution to this problem, so I don’t have to go thru all this hassle, in general I don’t know the answer in advance, so don’t have a clue about which corner, face, etc. to start with.)

(1) Try none of the constraints binding; KT says \( \bar{x}_0 = \bar{\lambda}_1 = \bar{\lambda}_2 = 0 \). which implies \( x = (-2, 2) \).

Contradiction. Assumed that \( \bar{x} \) was nonnegative; found that if there were a \( \bar{x} \) that satisfied
the KT conditions under these assumptions, then $\bar{x}_1$ would be negative. Note that we couldn’t have had a point satisfying this condition anyway, because of the non-vanishing gradient property, we checked above

(2) Try $\bar{x}_1 > 0$, $\bar{x}_2 > 0$ and $p \cdot \bar{x} = y$. KT conditions say that $\bar{\lambda}_1 = \bar{\lambda}_2 = 0$; $\bar{\lambda}_0 \geq 0$. Plugging these is gives

\begin{align*}
x_2 - 2 &= \lambda_0 p_1 \\
x_1 + 2 &= \lambda_0 p_2
\end{align*}

But when $p_1 = p_2$, this means that $x_2 - x_1 = 4$. On the other hand, since $p_1 = p_2 = y$, the budget constraint implies that $x_1 + x_2 = 1$. Substituting yields $2x_2 = 5$, which implies $x_1$ MUST be negative, contradicting our initial condition that $x_1$ must be nonnegative.

(3) Try $\bar{x}_1 > 0$, $\bar{x}_2 = 0$ and $p \cdot \bar{x} = y$. KT conditions say that $\bar{\lambda}_0, \bar{\lambda}_2 \geq 0$; $\bar{\lambda}_1 = 0$. Plugging these is gives

\begin{align*}
L_{\bar{x}_1} &= x_2 - 2 - \lambda_0 p_1 + \lambda_1 \\
&= -2 - \lambda_0 p_1 = 0;
\end{align*}

which implies $\lambda_0 = -2/p_1$, which is a contradiction.

(4) Try $\bar{x}_1 = 0$, $\bar{x}_2 > 0$ and $p \cdot \bar{x} = y$. KT conditions say that $\bar{\lambda}_0, \bar{\lambda}_1 \geq 0$; $\bar{\lambda}_2 = 0$. Plugging these in gives

\begin{align*}
L_{\lambda_0} &= y - p_2 x_2 = 0 \\
L_{\bar{x}_2} &= x_1 + 2 - \lambda_0 p_2 \\
&= 2 - \lambda_0 p_2 = 0
\end{align*}
which implies \( \lambda_0 = 2/p_2 \). Now consider \( L_{x_1} \), i.e.,

\[
L_{x_1} = x_2 - 2 - \lambda_0 p_1 + \lambda_1
\]

\[
= y/p_2 - 2 - 2 + \lambda_1 \quad \text{(since } \lambda_0 = 2/p_1 \text{)}
\]

\[
= 1 - 4 + \lambda_1
\]

which implies that \( \lambda_1 = 3 \). So we have a solution, i.e., \((0, y/p_2)\) with \( \lambda_0 = 2/p_2, \lambda_1 = 3 \).

4.8. **Computed solution to a NPP: ARE problem set example.**

This example was a homework problem for Econ 201A, 1999:

The example:

\[
\max u_i(x_i) = x_{1i}(4 - x_{2i}) \quad \text{s.t.}
\]

\[
p_1 x_{1i} + p_2 x_{2i} = p_1 \omega_{1i} + p_2 \omega_{2i};
\]

\[
x_i \geq 0;
\]

where \( \omega_1 = (4, 3), \omega_2 = (1, 0) \). The problem that 201 students faced was to solve for the \( x_i \)'s for \( i = 1, 2 \), and for the equilibrium prices. What I’ll do in these notes is to solve for the demand functions for good \#1, and to derive some equilibrium properties of the price vector.

For this problem, \( g \) is the matrix above, i.e.,

\[
\begin{bmatrix}
p_1 & p_2 \\
-1 & 0 \\
0 & -1
\end{bmatrix};
\]
while \( b \) is
\[
\begin{bmatrix}
p_1\omega_{1i} + p_2\omega_{2i} \\
0 \\
0
\end{bmatrix};
\]

We’ll normalize by setting \( p_1 = 1 \) and let \( \xi^i : \mathbb{R}^2 \to \mathbb{R}^2 \) denote the demand function for agent \( i \), i.e., the demand function is now a function only of relative prices. We know from the answer sheet that
\[
\xi^1(p_2) = \begin{cases} 
\left(\frac{4-p_2}{2}, \frac{4+7p_2}{2p_2}\right) & \text{if } p_2 \leq -4/7 \\
(4 + 3p_2, 0) & \text{otherwise}
\end{cases}
\]
\[
\xi^2(p_2) = \begin{cases} 
\left(\frac{1-4p_2}{2}, \frac{1+4p_2}{2p_2}\right) & \text{if } p_2 \leq -1/4 \\
(1, 0) & \text{otherwise}
\end{cases}
\]

We’ll now derive the demand functions for agent #1, and check that our answers agree with \( \xi^1 \). You should check as an exercise that you can repeat the same steps for agent #2, and arrive at \( \xi^2 \).

Before embarking on this problem, we’ll get some intuition for the solution. Note first from the definition of the utility function that

(1) provided that a positive quantity of good 1 is consumed, good 2 is a “bad”;

(2) as \( x_{2i} \) increases above 4, then good 1 becomes a bad also, and the gradient of utility is a negative vector.

(3) it may not be immediately obvious, but when \( x_{2i} > 4 \), \( u_i(\cdot) \) is a strictly quasi-convex function.
Figure 2. The problem facing agent #1

(4) now assume that both prices are both positive:

(a) if $x_{2i} < 4$, you cannot get an interior (i.e., strictly positive) solution to the KT conditions because the gradient of the constraint is a strictly positive vector while the gradient of the utility function has one positive and one negative component.

(b) if $x_{2i} > 4$, you can get an interior solution to the KT conditions because the gradient of the utility function is strictly negative. In this solution, the non-typical constraint will be binding, i.e., instead of wanting to move NE in the positive quadrant you want to move SW. However, in this case the utility function is quasi-convex not quasi-concave.

When you solve for an interior solution to the KT conditions, you’ll have found a minimum on the constraint set, not a maximum.

(5) conclude from this that you cannot obtain an interior maximum to this problem if both prices are positive.

Rather than exploring all the possibilities exhaustively, we’ll henceforth assume that $p_2 < 0$, i.e., good 2 is a bad. Now we’ll draw the picture. The gradient of the budget constraint points down
and to the right, i.e., SE., and the budget line is a positively sloped line through endowment point. Fig. 2 indicates the budget line with a relatively small negative price $p_2$; The optimum for this player is obviously a corner solution. Clearly, in order to get an interior solution to #1’s optimization problem, you have to flatten the budget line, i.e., lower $p_2$.

From now on, I’m going to dump all the $i$ subscripts since we’re only dealing with $i = 1$.

Set up the Lagrangian, setting $p_1 = 1$ and $y(p_2) = \omega_1 + p_2\omega_2$, i.e., $y_1(p_2) = 4 + 3p_2$ and $y_2(p_2) = 1$.

$$L(x, \lambda) = x_1(4 - x_2) + \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3(y(p_2) - x_1 - p_2x_2) + \lambda_4(x_1 + p_2x_2 - y(p_2))$$

Recall from (1), that the first order conditions were

$$\partial L(\bar{x}, \bar{\lambda})/\partial x_i = 0; \quad \partial L(\bar{x}, \bar{\lambda})/\partial \lambda_j \geq 0; \quad \bar{\lambda}_j \partial L(\bar{x}, \bar{\lambda})/\partial \lambda_j = 0. \quad (1)$$

In this particular problem, these conditions imply

$$L_{x_1} = (4 - x_2) + \lambda_1 - (\lambda_3 - \lambda_4) = 0$$

$$L_{x_2} = -x_1 + \lambda_2 - (\lambda_3 - \lambda_4)p_2 = 0$$

$$L_{\lambda_1} = x_1 \geq 0$$

$$L_{\lambda_2} = x_2 \geq 0$$

$$L_{\lambda_3} = y(p_2) - (x_1 + p_2x_2) \geq 0$$

$$L_{\lambda_4} = x_1 + p_2x_2 - y(p_2) \geq 0$$
Note that $L_{\lambda_3} > 0$ implies $L_{\lambda_4} < 0$ while $L_{\lambda_4} > 0$ implies $L_{\lambda_3} < 0$. Conclude that $L_{\lambda_3} = L_{\lambda_4} = 0$ leaving open the possibility that either $\lambda_3$ or $\lambda_4$ could be positive. Accordingly, it will be convenient to define $\lambda_0 = (\lambda_3 - \lambda_4)$, which can be either positive negative or zero. Thus for agent 1:

$$L_{x_1} = (4 - x_2) + \lambda_1 - \lambda_0 = 0$$
$$L_{x_2} = -x_1 + \lambda_2 - \lambda_0 p_2 = 0$$
$$L_{\lambda_0} = x_1 + p_2 x_2 - (4 + 3p_2) = 0$$
$$L_{\lambda_1} = x_1 \geq 0$$
$$L_{\lambda_2} = x_2 \geq 0$$

By inspection of the figure we can see that there are really two possibilities:

(A) $p_2 < 0$; the budget line alone is binding (if $p_2$ is large in abs value, i.e., budget line relatively flat). In this case,

$$\lambda_1 = \lambda_2 = 0, \lambda_0 = \lambda_3 > 0$$

(B) $p_2 < 0$; the budget line and the nonneg constraint on good 2 are both binding (if $p_2$ is small in abs value, i.e., budget line relatively steep). In this case,

$$\lambda_1 = 0, \lambda_2 > 0, \lambda_0 = \lambda_3 > 0$$

On the other hand, Fig. 2 suggests that there are several possibilities that we can exclude based on the Lagrangian conditions. We'll focus on one of them, just for practice, but there are many more that we won’t check.

(C) $p_2 > 0$ and $x_i > 0, i = 1, 2$, i.e., the budget line is the only constraint satisfied with equality.
We'll begin with (C), write down the Lagrangian system and show that all of the requirements cannot simultaneously be satisfied. From the mantra, we know the reason: the gradient of the budget line and the gradient of the objective have to be co-linear, but they can't be, because the objective's gradient points NE, while the budget's gradient points SE. Our task now is to show this using the Lagrangian. The conditions are:

\[
\begin{align*}
L_{x_1} &= (4 - x_2) - \lambda_0 = 0 \\
L_{x_2} &= - x_1 - \lambda_0 p_2 = 0 \\
L_{\lambda_0} &= x_1 + p_2 x_2 - (4 + 3p_2) = 0 \\
\lambda_1 &= 0; \\
\lambda_2 &= 0;
\end{align*}
\]

From \(L_{x_1}\) we have that

\[
\lambda_0 = 4 - x_2 \lambda_0 = -x_1/p_2 < 0
\]

so that, substituting into \(L_{x_1}\)

\[
(4 - x_2) + \frac{x_1}{p_2} = 0, \tag{2}
\]

or, since \(x_1 > 0\),

\[
(x_2 - 4) = \frac{x_1}{p_2} > 0
\]

But from \(L_{\lambda_0}\), we have that

\[
x_1 + p_2 x_2 = 4 + 3p_2 > 4
\]
or

\[ 4 - x_2 > \frac{x_1}{p_2} > 0 \]  \hspace{1cm} (3)

But, obviously, (2) and (3) cannot simultaneously be satisfied, so we’ve established that the set of conditions listed in (C) cannot hold. Note, moreover, that we obtained the contradiction by showing that if \( p_2 > 0 \), then the combination of \( L_{x_1} \) and \( L_{x_2} \) would then be inconsistent with \( L_{\lambda_0} \).

Now let’s consider the possibilities which from the figure, we know are possible. We will take each of possibilities (A) and (B) in turn, and see their implications for the Lagrangian system;

(A) the budget line alone is binding:

\[ L_{x_1} = (4 - x_2) - \lambda_0 = 0 \]
\[ L_{x_2} = -x_1 - \lambda_0 p_2 = 0 \]
\[ L_{\lambda_0} = x_1 + p_2 x_2 - (4 + 3p_2) = 0 \]
Solving this in the usual way:

(a) \(0 = (4 - x_2) + \frac{x_1}{p_2}\) (from \(L_{x_1}\))

(b) \(0 = p_2(4 - x_2) + x_1\) (rearranging (a))

\[= x_1 - p_2x_2 + 4p_2\]

(c) \((4 + 3p_2) = x_1 + p_2x_2\) (from \(L_{x_2}\))

(d) \((4 + 3p_2) = 2p_2x_2 - 4p_2\) (subtracting (b) from (c))

(e) \(x_2 = \frac{4 + 7p_2}{2p_2}\) (rearranging (d))

(f) \(x_1 = \frac{4 - p_2}{2}\) (subst (3) into (c))

Note that \(x_2 \geq 0\) iff \(|p_2| \leq 4/7\). Summarizing, (e) and (f) give us agent #1’s demand function for \(p_2 \in (-\infty, -4/7]\), i.e.,

\[\xi^1(p_2) = \left(\frac{4 + 7p_2}{2p_2}, \frac{4 - p_2}{2}\right)\]

(B) both budget line and nonneg constraint on 2 are binding:

\[L_{x_1} = 4 - \lambda_0 = 0\]

\[L_{x_2} = -x_1 + \lambda_2 - \lambda_0p_2 = 0\]

\[L_{\lambda_0} = x_1 - (4 + 3p_2) = 0\]

From \(L_{x_1} = 0, \lambda_0 = 4\). From \(L_{\lambda_0} = 0, x_1 = (4 + 3p_2)\). Plugging both values into \(L_{x_2} = 0\),

\[L_{x_2} = -(4 + 3p_2) + \lambda_2 - 4p_2\]

\[= -4 - 7p_2 + \lambda_2 = 0;\]
Now $L_{x_2}$ can be zero with $\lambda_2 \geq 0$ iff $-7p_2 - 4 \leq 0$, i.e., if $|p_2| \leq 4/7$. Therefore, we have now computed agent #1’s demand function for $p_2 \in (-4/7, 0]$, i.e.,

$$\xi^1(p_2) = (4 + 3p_2, 0)$$

4.9. Second Order conditions Without Quasi-Concavity

The right way to think about second order conditions for a constrained maximum is as follows: you have to ensure that there are no feasible changes in $x$ that will increase your objective function while keeping you in the constraint set. Divide the possible changes in $x$ that increase $f$ into

- changes that give you a first order increase in the objective function i.e., moves in a direction $dx$ such that $\nabla f(\bar{x})dx > 0$.
- changes that give you only a second order increase in the objective function i.e., moves in a direction $dx$ such that $\nabla f(\bar{x})dx \approx 0$, but which have a positive second term in the Taylor expansion. The key point to note is: for any given direction which is not orthogonal to the gradient vector, if the length of $dx$ is sufficiently small, then the first order term in the Taylor expansion dominates; however, for any $\epsilon > 0$, there will in general be directions that are nearly but not quite orthogonal to the gradient vector for which the first order term in the expansion will be dominated by the second. These are the directions that we take care of using the second order conditions.

You have to rule out the possibility of both kinds of changes, i.e., show that any such moves would take you outside of the constraint set.
• The K-T conditions do exactly the first of these: we saw that last time, i.e., we checked that if the K-T conditions were satisfied, then \textit{any} move in a direction \( \textbf{dx} \) such that \( \nabla f(\bar{x})\textbf{dx} > 0 \) took you outside of the constraint set.

• The K-T conditions have nothing to say about the second kind of change, and we have to rule them out, by looking at the second order Taylor expansion.

We’ll first consider second order conditions for maximizing subject to \textit{equality} constraints, and then see that we have really taken care of inequality constraints as well.

4.9.1. \textit{One linear equality constraint}. Consider the problem of maximizing a function subject to a single \textit{linear equality} constraint: maximize \( f(\textbf{x}) \) subject to \( g(\textbf{x}) = b \), where \( g \) is linear.

• we know that the first order condition for a maximum is that \( \nabla f(\textbf{x}) = \lambda \nabla g(\textbf{x}) \) for some \( \lambda \in \mathbb{R} \). Recall that this is only a necessary condition, not a sufficient one. Can’t distinguish between points \( a \) and \( c \) in Fig. 3:

• The second order condition for a maximum is that \( \textbf{dx}'Hf(\bar{x})\textbf{dx} < 0 \), for all \( \textbf{dx} \) such that \( \nabla g(\bar{x})\textbf{dx} = 0 \). i.e., you need to show that the “swivelling more than 90 degrees” condition holds only for vectors \( \textbf{dx} \) that lie in the particular \textit{subspace} of \( \mathbb{R}^n \) defined by the linear equality constraint, i.e., this is a much weaker condition. We say in this case that \textit{the Hessian is negative definite subject to the constraint} \( \nabla g(\bar{x})\textbf{dx} = 0 \).

• For some intuition for this condition, recall that to test for a local constrained max at \( \bar{x} \), we need to check that \( f(\textbf{x}) < f(\bar{x}) \), for all \( \textbf{x} \) that are in a nbd of \( \bar{x} \) and satisfy \( g(\textbf{x}) = b \). Now since \( g \) is linear, \( g(\textbf{x} + \textbf{dx}) = b \) if and only if \( \nabla g(\bar{x})\textbf{dx} = 0 \). So to test for a maximum, we can restrict our test to vectors that satisfy the condition \( \nabla g(\bar{x})\textbf{dx} = 0 \).
By Taylor’s theorem, we know that for vectors of this kind:

\[ f(\bar{x} + \delta x) - f(\bar{x}) = \nabla f(\bar{x}) \delta x + 0.5 \delta x \nabla^2 f(\bar{x}) \delta x + \text{Remainder} \]

\[ = \lambda \nabla g(\bar{x}) \delta x + 0.5 \delta x \nabla^2 f(\bar{x}) \delta x + \text{Remainder} \]

\[ = + 0.5 \delta x \nabla^2 f(\bar{x}) \delta x + \text{Remainder} \]

and for \( \delta x \) sufficiently small, the sign of the second order term determines the sign of the right hand side.

- If \( \delta x' \nabla^2 f(\bar{x}) \delta x < 0 \), for all \( \delta x \) such that \( \nabla g(\bar{x}) \delta x = 0 \), then the left hand side is necessarily negative.

- Thus, our second order condition, together with the first order condition, is necessary and sufficient for a constrained local max.
• How do you check to see if the above condition is satisfied? Answer: look at the bordered Hessian of $f$, just like we did when we checked for quasi-concavity. In this case, however, you border the Hessian of $f$ with the gradient of $g$.

• Fact: the condition

$$\mathbf{dx}'Hf(\bar{x})\mathbf{dx} < 0,$$
for all $\mathbf{dx}$ such that $\nabla g(\bar{x})\mathbf{dx} = 0$ (4)

holds if the sign of the $k$’th leading principal minor of the following bordered matrix has the same sign as $(-1)^k$:

$$\begin{bmatrix} 0 & g_1(\bar{x}) & g_2(\bar{x}) \\ g_1(\bar{x}) & f_{11}(\bar{x}) & f_{12}(\bar{x}) \\ g_2(\bar{x}) & f_{21}(\bar{x}) & f_{22}(\bar{x}) \end{bmatrix}.$$  (Recall that the $k$’th leading principal minor of this matrix is the determinant of the top-left $k+1 \times k+1$ submatrix.)

• Parenthetical Fact (for completeness): the condition

$$\mathbf{dx}'Hf(\bar{x})\mathbf{dx} > 0,$$
for all $\mathbf{dx}$ such that $\nabla g(\bar{x})\mathbf{dx} = 0$ (5)

holds if the sign of each leading principal minor of the above bordered matrix is negative.

• Of course, an alternative way of proceeding would have been to check that $f$ was quasi-concave, i.e., to have checked the bordered Hessian of $f$.

• Note that the test above is practically identical to the test for quasi-concavity of $f$. I.e., recall that to check for quasi-concavity of $f$, we look at the minors of the matrix

$$\begin{bmatrix} 0 & f_1(\bar{x}) & f_2(\bar{x}) \\ f_1(\bar{x}) & f_{11}(\bar{x}) & f_{12}(\bar{x}) \\ f_2(\bar{x}) & f_{21}(\bar{x}) & f_{22}(\bar{x}) \end{bmatrix}.$$  What’s the difference between these two bordered Hessian? In the first case, the border is the gradient of $g$ at $\bar{x}$; in the second it is the gradient of $f$ at $\bar{x}$. However, at the solution to the optimization problem, $\bar{x}$, we have $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$, so that for each $k$, the $k$’th principal minor of each of the two bordered matrices have the same signs. (I.e., the constant doesn’t affect the signs of the determinants.)
• Alternative, could think about a quasi-concave function as having the property of a concave function, provided we restrict attention to the subspace defined by the level set, i.e., \( Hf(x) \) is negative definite on the subspace \( \{ dx : \nabla f(x)dx = 0 \} \).

• Difference between checking for quasi-concavity and checking definiteness subject to constraint is that in the latter case, you only have to check the matrix for a specific value in the domain, whereas in the former you have to check for all possible values in the domain.

• When checking for quasi-concavity of \( f \) at \( x \), you check that \( f \) goes down as you move along the tangent plane to the level set of \( f \) through \( x \). Moreover, at \( \bar{x} \), this tangent plane is also the linear constraint \( g(\cdot) = b \).

4.9.2. One nonlinear equality constraint. Now consider the problem of maximizing a function subject to a single nonlinear equality constraint: maximize \( f(x) \) subject to \( g(x) = b \), where \( g \) is now nonlinear. Note that in this case, the cookie cutter sufficiency conditions cannot ever be satisfied:

• you have to maximize \( f \) subject to being in the lower contour set of \( g \) corresponding to \( b \) and subject to being in the upper contour set of \( g \) corresponding to \( b \). Because \( g \) is not affine (constraint is not linear), if the lower contour set of \( g \) is convex, the upper contour set won’t be.

• if the upper contour set of \( g \) is convex, the lower contour set won’t be.

So what do we do? Consider for example the problem set question, maximize or minimize the distance from a given ellipse to the origin.

• if we are maximizing, we need the surface of the ellipse to be more curved than the level set of the distance function (i.e., more curved than a circle)
• if we are minimizing, we need the surface of the ellipse to be less curved than the level set of the distance function (i.e., less curved than a circle)

Fortunately, unless the ellipse is itself a circle, then it will be flatter than a circle at its flattest point and more curved than a circle at its most curved point.

Looks like we only have a local solution. Check this.

In this case, we need to consider the Hessian of the Lagrangian, evaluated at the point \((\bar{x}, \bar{\lambda})\) satisfying its first order conditions, rather than simply the Hessian of the objective function. That is, instead of looking at the bordered matrix:

\[
B\mathbf{H}f(\bar{x}) = \begin{bmatrix}
0 & g_1(\bar{x}) & g_2(\bar{x}) \\
g_1(\bar{x}) & f_{11}(\bar{x}) & f_{12}(\bar{x}) \\
g_2(\bar{x}) & f_{21}(\bar{x}) & f_{22}(\bar{x})
\end{bmatrix}
\]

you look at the following matrix, which is the (unbordered) Hessian of the Lagrangian:

\[
\mathbf{H}L(\bar{x}) = \begin{bmatrix}
0 & g_1(\bar{x}) & g_2(\bar{x}) \\
g_1(\bar{x}) & f_{11}(\bar{x}) - \bar{\lambda}g_{11}(\bar{x}) & f_{12}(\bar{x}) - \bar{\lambda}g_{12}(\bar{x}) \\
g_2(\bar{x}) & f_{21}(\bar{x}) - \bar{\lambda}g_{21}(\bar{x}) & f_{22}(\bar{x}) - \bar{\lambda}g_{22}(\bar{x})
\end{bmatrix}
\] (6)

For the case of one nonlinear inequality constraint, the second order condition for a maximum is that

\[
\mathbf{d}x' (\mathbf{H}f(\bar{x}) - \bar{\lambda}\mathbf{H}g(\bar{x})) \mathbf{d}x < 0, \text{ for all } \mathbf{d}x \text{ such that } \nabla g(\bar{x})\mathbf{d}x = 0,
\] (7)
or, equivalently, that the \(k\)'th principal minor of the matrix \(\mathbf{H}L(\bar{x})\) has the same sign as \((-1)^k\).

(Even though \(\mathbf{H}L(\bar{x})\) is technically an unbordered hessian, when we talk about its minors, it may
Figure 4. Role of Second Order Condition for a Constrained Max

as well be, i.e., the first minor is the determinant of 

\[
\begin{bmatrix}
0 & g_1(\bar{x}) \\
g_1(\bar{x}) & f_{11}(\bar{x}) - \bar{\lambda}g_{11}(\bar{x})
\end{bmatrix},
\text{etc.}
\]

Why the difference in the nonlinear case? I.e., why doesn’t the \( \lambda \) term show up in the linear case?

- First note that in the linear case, \( \text{HL}(\bar{x}) \) reduces to \( \text{BHf}(\bar{x}) \) since the second derivatives of \( g \) are all zero.
- Graphically, note that even if \( f \) had the right curvature (i.e., if the upper contour set of \( f \) through \( \bar{x} \) were a convex set) the point that the KT conditions has located could be a minimum not a maximum because the \( g \) function could have the wrong curvature (see the right panel of Fig. 4). But unless you assume that \( g \) is quasi-convex, you can’t rule out the possibility that the level set of \( g \) is also convex to the origin, and has greater curvature than \( f \). In this case, your KT conditions would indeed locate a min.
- what determines the curvature of \( g \)? Two things:
  - the Hessian of \( g \);
  - the gradient of \( g \).
  - to see the relationship, suppose that \( g \) is negative definite at \( \bar{x} \) (not positive definite as usual), pick a direction \( \text{dx} \) that makes an angle with \( \nabla g(\bar{x}) \) that is acute but close to 90deg. Consider (a) the level set of \( g \) passing thru \( \bar{x} \) and (b) the line starting at
Figure 5. Relationship between Curvature of $g$ and length of $\nabla g$

$\bar{x}$ that points in the direction $dx$. Finally, identify the closest point to $\bar{x}$ on the line (b) that intersects the level set (a). (Such a point must exist if $g$ is negative definite at $\bar{x}$ and the angle between $dx$ and $\nabla g(\bar{x})$ is sufficiently close to 90deg.) See Fig. 5.

Specifically, solve for the point $dx^*$ that satisfies

$$g(\bar{x}) + dx^* - g(\bar{x}) = 0 = \nabla g(\bar{x})dx^* + 0.5dx^*H_g(\bar{x})dx^* + \text{Remainder} \quad (8)$$

- Observe that if in equation 8, you double the length of $\nabla g$ holding everything else constant, you’d have to lengthen $dx^*$ to increase the relative importance of the second term and restore the equality. In other words, increasing $\nabla g$ holding all second partials constant decreases the curvature of the level set.

It’s for this reason that in the expression for the second order condition—i.e., display (6)—the second partials of $g$ are multiplied by $\lambda$:

- recall that $\lambda = \frac{||\nabla f(\bar{x})||}{||\nabla g(\bar{x})||}$. 
therefore, holding \(|\nabla f(\bar{x})|\) constant, the smaller is \(|\nabla g(\bar{x})|\), the higher is \(\lambda\), i.e., more curved is the constraint set, i.e., the more likely you are to be in the case illustrated by the right hand panel of Fig. 4 (where \(\bar{x}\) is not even a local max) than the left hand panel (where \(\bar{x}\) is a local max).

• Thus, in expression (6) for the second order conditions above, the role of \(\lambda\) should now be clear: if as in the right panel of Fig. 4, \(f\) is strictly quasi concave and \(g\) is locally negative definite, then holding everything else constant, as \(|\nabla g(\bar{x})|\) decreases, \(\lambda\) increases, the curvature of the constraint set increases and it becomes harder to satisfy the second order condition.

To illustrate further, suppose for the moment that the cross partials of \(f\) and \(g\) are both zero. In this case, we have

\[
\text{HL}(\bar{x}) = \begin{bmatrix}
0 & g_1(\bar{x}) & g_2(\bar{x}) \\
g_1(\bar{x}) & f_{11}(\bar{x}) - \lambda g_{11}(\bar{x}) & 0 \\
g_2(\bar{x}) & 0 & f_{22}(\bar{x}) - \lambda g_{22}(\bar{x})
\end{bmatrix}
\]

multiplying the diagonals, we get

\[
\text{det} \left( \text{HL}(\bar{x}) \right) = - \left\{ g_2(\bar{x})^2 (f_{11}(\bar{x}) - \lambda g_{11}(\bar{x})) + g_1(\bar{x})^2 (f_{22}(\bar{x}) - \lambda g_{22}(\bar{x})) \right\}
\]

which is positive if \(f_{22}(\bar{x}) < \lambda g_{22}(\bar{x})\) and \(f_{11}(\bar{x}) < \lambda g_{11}(\bar{x})\). This is certainly true if \(f_{11}, f_{22}\) are negative (quasi-concavity) while \(g_{11}, g_{22}\) are positive (quasi-convexity). On the other hand it is also true if for \(i = 1, 2\), \(f_{ii}\) is larger in absolute magnitude than \(\lambda g_{ii}\).