ECON 186 Class Notes: Optimization Part 3

Jijian Fan

Second Order Conditions for Constrained Optimization

- Suppose we have a function $z = f(x, y)$ subject to $g(x, y) = c$. The second order conditions in the constrained case still revolve around positive and negative definiteness, but instead of being concerned with all possible values of dx and dy, we want only the values of dx and dy that satisfy the linear constraint

$$g_x dx + g_y dy = 0$$

- Then the second-order necessary conditions are:
  - For maximum of z: d^2z negative semidefinite, subject to $dg = 0$
  - For minimum of z: d^2z positive semidefinite, subject to $dg = 0$
- The second-order sufficient order conditions are:
  - For maximum of z: d^2z negative definite, subject to $dg = 0$
  - For minimum of z: d^2z positive definite, subject to $dg = 0$

Second Order Conditions for Constrained Optimization

- Suppose that we are a given a function $f(x, y)$ then we recall that the Hessian of f is

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}$$

- However, again consider the case where $f(x, y)$ now has a constraint so that the Lagrangian can be written as

$$L = f(x, y) + \lambda [c - g(x, y)]$$

- In the constrained case, we determine positive and negative definiteness by the bordered Hessian, which is simply the Hessian of the Lagrangian function "bordered" by the first derivatives of the constraint. So, the conditions are
- $d^2 z$ is $\left\{ \begin{matrix} \textit{positive definite} \\ \textit{negative definite} \end{matrix} \right\}$ subject to $dg = 0$ iff

$$\begin{vmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{vmatrix} \begin{cases} < & 0 \\ > & 0 \end{cases}$$

Second Order Conditions - Multiconstraint Case

- Suppose we are attempting to optimize a function with n constraints. $Z = f(x_1, ..., x_n) + \sum_{j=1}^{m} \lambda_j \left[c_j - g^j(x_1, ..., x_n)\right]$
- Then, the bordered Hessian is

$$|\overline{H}| = \begin{vmatrix} 0 & 0 & \cdots & 0 & \vdots & g_1^1 & g_2^1 & \cdots & g_n^1 \\ 0 & 0 & \cdots & 0 & \vdots & g_1^2 & g_2^2 & \cdots & g_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \vdots & g_1^m & g_2^m & \cdots & g_n^m \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ g_1^1 & g_1^2 & \cdots & g_1^m & \vdots & Z_{11} & Z_{12} & \cdots & Z_{1n} \\ g_2^1 & g_2^2 & \cdots & g_2^m & \vdots & Z_{21} & Z_{22} & \cdots & Z_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ g_n^1 & g_n^2 & \cdots & g_n^m & \vdots & Z_{n1} & Z_{n2} & \cdots & Z_{nn} \end{vmatrix}$$

- For a maximum of z, a sufficient condition is that $|\overline{H}_{m+1}|, |\overline{H}_{m+2}|, ..., |\overline{H}_n|$ alternates in sign, where $|\overline{H}_{m+1}|$ has sign $(-1)^{m+1}$.
- For a minimum of z, a sufficient condition is that $|\overline{H}_{m+1}|, |\overline{H}_{m+2}|, ..., |\overline{H}_n|$ all have sign $(-1)^m$.

Lagrangian Example

- A politician facing reelection can win votes according to the following process: $c = 500S^{0.2}M^{0.6}$, where S is hours of making campaign speeches and M is the number of flyers mailed. Making speeches costs \$10 per hour, mailing flyers costs \$.50 per flyer, and \$8,000 are available to spend on the campaign. Assuming the politician wants to maximize votes, how should the budget be allocated between speeches and mailing flyers?

$$U = 500S^{0.2}M^{0.6}$$

$$Budget\ constraint : p_s s + p_m m = y = g(s,m) \rightarrow 10s + .5m = 8000$$

$$L = 500S^{0.2}M^{0.6} + \lambda\left(8000 - 10S - 0.5M\right)$$

Lagrangian Example

- First order conditions:

$$\frac{\partial L}{\partial S} = 0.2(500)S^{-0.8}M^{0.6} - 10\lambda = 0 \rightarrow \lambda = \frac{0.2(500)M^{0.6}}{10S^{0.8}}$$

$$\frac{\partial L}{\partial M} = 0.6(500)S^{0.2}M^{-0.4} - .5\lambda = 0 \rightarrow \lambda = \frac{0.6(500)S^{0.2}}{0.5M^{0.4}}$$

$$\frac{\partial L}{\partial \lambda} = 8000 - 10S - 0.5M = 0$$

Example

$$\frac{0.2(500)M^{0.6}}{10S^{0.8}} = \frac{0.6(500)S^{0.2}}{0.5M^{0.4}} \rightarrow \frac{10M^{0.6}}{S^{0.8}} = \frac{600S^{0.2}}{M^{0.4}} \rightarrow M = 60S$$

$$10S + .5(60S) = 8000 \rightarrow 40S = 8000 \rightarrow S^* = 200$$

$$M^* = 60(200) = 12000$$

- So, the politician can receive the most votes by spending \$10(200) = \$2000 on speeches and \$0.5(12000) = \$6000 on flyers.

Example

- Second-order conditions:

$$L_{ss} = 100(-0.8)S^{-1.8}M^{0.6} = -80S^{-1.8}M^{0.6}$$

$$L_{mm} = 300(0.2)S^{0.2}(-0.4)M^{-1.4} = -24S^{0.2}M^{-1.4}$$

$$L_{sm} = 100(0.6)S^{-0.8}M^{-0.4} = 60S^{-0.8}M^{-0.4}$$

$$g_s = 10, \; g_y = .5$$

Example

- So the bordered Hessian is

$$\begin{vmatrix} 0 & 10 & .5 \\ 10 & -80 & 60 \\ .5 & 60 & -24 \end{vmatrix} = -10 \begin{vmatrix} 10 & 60 \\ .5 & -24 \end{vmatrix} + .5 \begin{vmatrix} 10 & -80 \\ .5 & 60 \end{vmatrix} =$$

$$= -10(-240-30) + .5(600+40) = 2700 + 320 = 3020 > 0$$

- So, the bordered Hessian is negative definite, which means that the optimal values we found previously are in fact a maximum.

Multiconstraint Second Order Condition Example

- Suppose we want to find the critical points of the function $f(x,y,z) = z$ subject to $g_1(x,y,z) = x + y + z = 12$ and $g_2(x,y,z) = x^2 + y^2 - z = 0$. Then the Lagrangian function is

$$L(x,y,z) = z + \lambda(12 - x - y - z) + \mu\left(-x^2 - y^2 + z\right)$$

- We can form the bordered Hessian

$$|H_3| = \begin{vmatrix} 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & -2x & 2y & 1 \\ -1 & -2x & -2\mu & 0 & 0 \\ -1 & -2y & 0 & -2\mu & 0 \\ -1 & 1 & 0 & 0 & 0 \end{vmatrix}$$

- Since this determinant is very complicated to compute, we will refrain from computing it.

Homogenous Functions

- A function is homogeneous of degree r if multiplication of each of its independent variables by a constant j will alter the value of the function by the proportion j^r, that is, if

$$f\left(jx_1, \dots, jx_n\right) = j^r f\left(x_1, \dots, x_n\right)$$

- Example 1: Consider the function $f(x, y, w) = \frac{x}{y} + \frac{2w}{3x}$. Then, if we multiply each variable by j, we get

$$f(jx, jy, jw) = \frac{(jx)}{(jy)} + \frac{2(jw)}{3(jx)} = \frac{x}{y} + \frac{2w}{3x} = f(x, y, w) = j^0 f(x, y, w)$$

  - So, f is a homogeneous function of degree zero.

Homogeneous Functions

- Example 2: Consider the function $g(x,y,w)=\frac{x^2}{y}+\frac{2w^2}{x}$. Then, multiplying through by j, we get

$$g(jx,jy,jw)=\frac{(jx)^2}{(jy)}+\frac{2\,(jw)^2}{(jx)}=j\left(\frac{x^2}{y}+\frac{2w^2}{x}\right)=jg(x,y,w)$$

- So, g is homogeneous of degree one. That is, multiplication of each variable by j will alter the value of the function exactly j-fold.
- Example 3: Consider the function $h(x,y,w)=2x^2+3yw-w^2$. Multiplying each variable by j will give us

$$h(jx,jy,jw)=2\,(jx)^2+3\,(jy)\,(jw)-(jw)^2=j^2h\,(x,y,w)$$

- So, h is homogeneous of degree two. For example, doubling each variable will quadruple the value of the function.

Economic Application of Homogeneous Functions

- Functions that are homogeneous of degree one are also known as linearly homogeneous functions (note that they do not have to be linear though!), and the main economic application is to that of production functions.
- Consider a production function where quantity depends on capital (K) and labor (L). That is,

$$Q = f(K, L)$$

Economic Application of Homogeneous Functions

- Assume that $Q = f(K, L)$ is linearly homogeneous, that is, homogeneous of degree one. Then, there are three useful properties of $Q = f(K, L)$.
  - Property 1: The average product of labor (AP_L) and of capital (AP_K) can be expressed as functions of the capital-labor ratio, $k \equiv \frac{K}{L}$, alone and are homogeneous of degree zero.
  - Property 2: The marginal product of labor (MP_L) and of capital (MP_K) can be expressed as functions of k alone and are homogeneous of degree zero.
  - Property 3: $K\frac{\partial Q}{\partial K} + L\frac{\partial Q}{\partial L} \equiv Q$
- These properties represent why a linearly homogeneous production function is said to have constant returns to scale.

Proof of Property 1 of Linearly Homogeneous Production Function

- Proof of Property 1
  - Multiply each independent variable, that is, K and L by $\frac{1}{L}$. We get

$$\frac{Q}{L} = f\left(\frac{K}{L}, \frac{L}{L}\right) = f\left(\frac{K}{L}, 1\right) = f(k, 1)$$

  - Since K and L are to be replaced with k and 1, respectively, each time they appear, f is now simply a function of k, which we call $\phi(k)$. So,

$$\frac{Q}{L} = \phi(k)$$

  - However, we know that the average products are simply total quantity divided by the amount of the respective input, so

$$AP_L \equiv \frac{Q}{L} = \phi(k)$$

$$MP_L \equiv \frac{Q}{K} = \frac{Q}{L}\frac{L}{K} = \frac{\phi(k)}{k}$$

Cobb-Douglas Production Function

- A very widely used production function is called the Cobb-Douglas production function and takes the general form:

$$Q = AK^{\alpha}L^{\beta}$$

- To check the homogeneity of this production function, multiply each input by j to get

$$A(jK)^{\alpha}(jL)^{\beta} = j^{\alpha+\beta}\left(AK^{\alpha}L^{\beta}\right) = j^{\alpha+\beta}Q$$

- So, Q is homogeneous of degree $(\alpha+\beta)$.
- Therefore, the Cobb-Douglas production function is said to have constant returns to scale if $\alpha+\beta=1$, since it is linearly homogeneous.
- Additionally, if $\alpha+\beta<1$, the function is said to have decreasing returns to scale, and if $\alpha+\beta>1$, the function has increasing returns to scale.

Nonlinear Programming and Kuhn-Tucker Conditions

- So far, we have analyzed only linear and binding (equality) constraints. However, optimization methodology extends naturally to nonlinear constraints and objective functions, as well as inequality constraints.
- With no inequality constraints or sign restrictions on the choice variables, the first-order condition for a relative extremum is simply that the first partial derivatives of the Lagrangian function with respect to all Lagrange multipliers and choice variables be zero.
    - However, in nonlinear programming, similar conditions are called Kuhn-Tucker conditions.

Nonlinear Programming and Kuhn-Tucker Conditions

- First, consider the single-variable case where x_1 is restricted to be nonnegative. That is, we want to maximize $\pi = f(x_1)$ subject to $x_1 \geq 0$.
  - Three situations arise as in Figure 13.1 of Chaing and Wainwright, page 403, 4th edition:
    - ⋆ (a) A relative maximum of y occurs in the interior of the shaded feasible area, such as point A in Figure 13.1, then we have an interior solution. The FOC in this case is $\frac{dy}{dx} = f^{'}(x) = 0$, same as usual.
    - ⋆ (b) A relative maximum can also occur on the vertical axis, shown in point B, where x=0. Even here, where we have a boundary solution, the FOC $\frac{dy}{dx} = f^{'}(x) = 0$ is still valid.
    - ⋆ (c) A relative maximum may take the position of points C or D, because to qualify as a local maximum the point simply has to be higher than the neighbouring points within the feasible region. Here, $f^{'}(x) < 0$.

Nonlinear Programming and Kuhn-Tucker Conditions

- We can combine these statements into a single set of statements: $f^{'}(x_1) \leq 0$, $x_1 \geq 0$, and $x_1 f^{'}(x_1) = 0$.
- The third statement is called the complementary slackness condition because either x_1 or $f^{'}(x_1)$ must equal 0.
- Generalizing the problem to n variables, the problem becomes: maximize $\pi = f(x_1, x_2, ..., x_n)$ subject to $x_j \geq 0 \qquad (j = 1, 2, ..., n)$
- Then, the Kuhn-Tucker conditions become

$$f_j \leq 0, \qquad x_j \geq 0, \qquad x_j f_j = 0 \qquad (j = 1, 2, ..., n)$$

General Inequality Constraints

- Suppose we want to optimize a function $f(x_1, x_2, ..., x_n)$ subject to m constraints of the form $g^i(x_1, x_2, ..., x_n) = r_i \qquad (i = 1, ..., m)$ where the $g^i s$ are inequality constraints.
- Then, we can write the Lagrangian function as

$$L = f(x_1, x_2, ..., x_n) + \sum_{i=1}^{m} \lambda_i \left[r_i - g^i(x_1, x_2, ..., x_n) \right]$$

- The Kuhn-Tucker conditions are:

$$L_{x_j} \leq 0 \qquad x_j \geq 0 \qquad \textit{and} \qquad x_j L_{x_j} = 0 \qquad [\textit{maximization}]$$

$$L_{\lambda_i} \geq 0 \qquad \lambda_i \geq 0 \qquad \textit{and} \qquad \lambda_i L_{\lambda_i} = 0 \qquad \begin{pmatrix} i = 1, 2, ..., m \\ j = 1, 2, ..., n \end{pmatrix}$$

Nonlinear Programming Example

- Maximize $U = xy$ subject to $x + y \leq 100$ and $x \leq 40$ and $x, y \geq 0$.
- The Lagrangian function is

$$L = xy + \lambda_1 (100 - x - y) + \lambda_2 (40 - x)$$

- The Kuhn-Tucker conditions are:

$$L_x = y - \lambda_1 - \lambda_2 \leq 0 \qquad x \geq 0 \qquad \text{and} \qquad xL_x = 0 \tag{1}$$

$$L_y = x - \lambda_1 \leq 0 \qquad y \geq 0 \qquad \text{and} \qquad yL_y = 0 \tag{2}$$

$$L_{\lambda_1} = 100 - x - y \geq 0 \qquad \lambda_1 \geq 0 \qquad \text{and} \qquad \lambda_1 L_{\lambda_1} = 0 \tag{3}$$

$$L_{\lambda_2} = 40 - x \geq 0 \qquad \lambda_2 \geq 0 \qquad \text{and} \qquad \lambda_2 L_{\lambda_2} = 0 \tag{4}$$

Nonlinear Programming Example

- We solve nonlinear programming problems by trial and error. Specifically, we look at each case and see if the solution violates the inequality constraints.
- First, consider $x = 0, y = 0$ or $x = 0, y > 0$ or $x > 0, y = 0$. In any of these cases, $\underline{U = 0}$ which cannot possibly be the maximum since $x \geq 0$ and $y \geq 0$.
- Then, we consider only where $x > 0$ and $y > 0$ (so $L_x = L_y = 0$), and we have 4 cases from the complementary slackness conditions for the constraints.
- Case 1: $\lambda_1 > 0, \lambda_2 > 0, x > 0, y > 0$
  - By complementary slackness, $100 - x - y = 0$ and $40 - x = 0$. Combining the two, we get $x^* = 40$ and $y^* = 100 - 40 = 60$. Then, (1) and (2) give $60 - \lambda_1 - \lambda_2 = 0$ and $40 - \lambda_1 = 0$. So, $\lambda_1 = 40$ and $\lambda_2 = 60 - 40 = 20$. This satisfies all of the constraints so this is a solution.

Nonlinear Programming Example

- Case 2: $\lambda_1 > 0, \lambda_2 = 0, x > 0, y > 0$
  - In this case, from complementary slackness, $100 - x - y = 0$. So, $x = 100 - y$. Then, plugging into (2), we get $\lambda_1 = 100 - y$. Plugging into (1), we get $y - (100 - y) = 0$, so $y^* = 50$ and therefore $x^* = 50$. However, this violates the constraint $x \leq 40$. So, this cannot be a solution.
- Case 3: $\lambda_1 = 0, \lambda_2 > 0, x > 0, y > 0$
  - From complementary slackness, we know $x^* = 40$. Then, from (2), $\lambda_1^* = x^* = 40$. However, this is not possible since $\lambda_1 = 0$, so this cannot be a solution.
- Case 4: $\lambda_1 = 0, \lambda_2 = 0, x > 0, y > 0$
  - From (1) and (2), we would get $x^* = y^* = 0$, but this is a contradiction since we are assuming $x > 0, y > 0$.

Minimization

- Suppose we instead want to minimize a function with inequality constraints, then we use the fact that minimizing C is equivalent to maximizing $-C$, but we must also flip the inequality constraints.
- So, suppose that we want to minimize a function f where the Lagrangian function is $L = f(x_1, x_2, ..., x_n) + \sum_{i=1}^{m} \lambda_i \left[r_i - g^i(x_1, x_2, ..., x_n) \right]$ and $x_i \geq 0$ for $i = 1, 2, ..., n$.
- Then, the Kuhn-Tucker conditions for minimization are:

$$\frac{\partial L}{\partial x_j} \geq 0 \qquad x_j \geq 0 \qquad x_j \frac{\partial L}{\partial x_j} = 0 \qquad [minimization]$$

$$\frac{\partial L}{\partial \lambda_j} \leq 0 \qquad \lambda_i \geq 0 \qquad \lambda_i \frac{\partial L}{\partial \lambda_i} = 0 \qquad \begin{pmatrix} i = 1, 2, ..., m \\ j = 1, 2, ..., n \end{pmatrix}$$

Nonlinear Programming Minimization Example

- Minimize $C = (x_1 - 4)^2 + (x_2 - 4)^2$ subject to $2x_1 + 3x_2 \geq 6$ and $-3x_1 - 2x_2 \geq -12$ and $x_1, x_2 \geq 0$.

$$L = (x_1 - 4)^2 + (x_2 - 4)^2 + \lambda_1 (6 - 2x_1 - 3x_2) + \lambda_2 (-12 + 3x_1 + 2x_2)$$

- The Kuhn-Tucker conditions are:

$$L_{\lambda_1} : 6 - 2x_1 - 3x_2 \leq 0 \qquad \lambda_1 \geq 0 \qquad \lambda_1 L_{\lambda_1} = 0$$

$$L_{\lambda_2} : -12 + 3x_1 + 2x_2 \leq 0 \qquad \lambda_2 \geq 0 \qquad \lambda_2 L_{\lambda_2} = 0$$

$$L_{x_1} : 2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 \geq 0 \qquad x_1 \geq 0 \qquad x_1 L_{x_1}$$

$$L_{x_2} : 2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 \geq 0 \qquad x_2 \geq 0 \qquad x_2 L_{x_2}$$

Nonlinear Programming Minimization Example

- First, let's look at the cases for x_1 and x_2:
- Case 1: $x_1 = 0$, $x_2 = 0$
  - This does not satisfy the constraint $2x_1 + 3x_2 \geq 6$.
- Case 2: $x_1 = 0, x_2 > 0$
  - Using the two constraints, we get that $x_2 \in [2, 6]$. Plugging in $x_2 = 4$ to C, we get $C = (x_1 - 4)^2 + (x_2 - 4)^2 = 16 + 0 = 16$. However, take for example $x_1 = 2, x_2 = 2$ which satisfies the constraints and gives us $C = (2-4)^2 + (2-4)^2 = 4 + 4 = 8$. So there is no way $x_1 = 0, x_2 > 0$ could provide a solution that minimizes C.
- Case 3: $x_1 > 0, x_2 = 0$
  - Similiarly to Case 2, $x_1 \in [3, 4]$, so plugging in 4, we get $C = 0 + 16 = 16$. Again, we could easily find a smaller value when $x_1 > 0, x_2 > 0$.

Nonlinear Programming Minimization Example

- So, it must be the case that $x_1 > 0, x_2 > 0$. Then we must look at the 4 cases for the $\lambda_i^{'}s$.
- Case 1: $\lambda_1 > 0, \lambda_2 > 0, x_1 > 0, x_2 > 0$
  - Then, by complementary slackness, $6 - 2x_1 - 3x_2 = 0$ and $-12 + 3x_1 + 2x_2 = 0$. Solving the first for x_1, we get $x_1 = \frac{6-3x_2}{2}$. Plugging into the second, $-12 + 3\left(\frac{6-3x_2}{2}\right) + 2x_2 = 0 \rightarrow -12 + 9 - \frac{9}{2}x_2 + 2x_2 = 0 \rightarrow -3 = \frac{5}{2}x_2 \rightarrow x_2 = -\frac{6}{5}$. This violates the nonnegativity constraint.
- Case 2: $\lambda_1 > 0, \lambda_2 = 0, x_1 > 0, x_2 > 0$
  - By complementary slackness, $6 - 2x_1 - 3x_2 = 0$. From the FOC for x_1, $2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 = 2(x_1 - 4) - 2\lambda_1 = 0 \rightarrow \lambda_1 = x_1 - 4$. From the FOC for x_2, $2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 = 2(x_2 - 4) - 3\lambda_1 = 0 \rightarrow \lambda_1 = \frac{2(x_2-4)}{3}$. So, $\lambda_1 = \lambda_1 \rightarrow x_1 - 4 = \frac{2(x_2-4)}{3} \rightarrow x_1 = \frac{2x_2}{3} + \frac{4}{3} = \frac{2}{3}(x_2 + 2)$. Plugging into the constraint, we get $6 - 2\left(\frac{2}{3}(x_2+2)\right) - 3x_2 = 0 \rightarrow 6 - \frac{4}{3}x_2 - \frac{8}{3} - 3x_2 = 0 \rightarrow \frac{10}{3} - \frac{13}{3}x_2 = 0 \rightarrow x_2^* = \frac{10}{13}$. Now, we can see that if we plug in to the equation for λ_1, we get $\lambda_1 = \frac{2}{3}(x_2 - 4) = \frac{2}{3}\left(\frac{10}{13} - 4\right) < 0$, so this cannot be a solution.

Nonlinear Programming Minimization Example

- Case 3: $\lambda_1 = 0, \lambda_2 > 0, x_1 > 0, x_2 > 0$
  - From the FOC for x_1,
    $2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 = 0 \rightarrow -3\lambda_2 = 2(x_1 - 4) \rightarrow \lambda_2 = -\frac{2}{3}(x_1 - 4)$.
    From the FOC for x_2,
    $2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 = 0 \rightarrow -2\lambda_2 = 2(x_2 - 4) \rightarrow \lambda_2 = 4 - x_2$. Then,
    $\lambda_2 = \lambda_2 \rightarrow -\frac{2}{3}(x_1 - 4) = 4 - x_2 \rightarrow x_2 = \frac{2}{3}x_1 + \frac{4}{3} = \frac{2}{3}(x_1 + 2)$. From complementary slackness, $-12 + 3x_1 + 2x_2 = 0$. Plugging in,
    $-12 + 3x_1 + 2\left(\frac{2}{3}(x_1 + 2)\right) = 0 \rightarrow -12 + 3x_1 + \frac{4}{3}x_1 + \frac{8}{3} = 0 \rightarrow$
    $\frac{13}{3}x_1 - \frac{28}{3} = 0 \rightarrow x_1^* = \frac{28}{13}$. Plugging back in,
    $x_2^* = \frac{2}{3}\left(\frac{28}{13} + 2\right) = \frac{2}{3}\left(\frac{54}{13}\right) = \frac{36}{13}$. This satisfies all the constraints, so it is a solution.

- Case 4: $\lambda_1 = 0, \lambda_2 = 0, x_1 > 0, x_2 > 0$
  - From the FOC for x_1, $2(x_1 - 4) = 0 \rightarrow x_1^* = 4$. From the FOC for x_2, $2(x_2 - 4) = 0 \rightarrow x_2^* = 4$. This violates the constraint $-3x_1 - 2x_2 \geq -12$ because $-3(4) - 2(4) = -20 \leq -12$, so this is not a solution.