ECON 501 Macroeconomic Analysis

Prof. Marco Cozzi

Notes on Multivariate Calculus and Optimization

1 Multivariate Calculus

For didactical purposes we restrict attention to real-valued functions of two exogenous variables. Everything applies readily to real-valued functions of n exogenous variables.

Definition: a function of two variables is a mapping $f : \Re_2 \longrightarrow \Re$, $\vec{x} \mapsto f(x) = z$, where $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$. Therefore we often write

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

Note that continuity is defined as in the single-variable case, only that convergence of $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ towards some point $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ now happens component-wise.

Often we want to determine the slope of f at some point (x_0, y_0). However, the graph of a function of two exogenous variables is in three dimensions and so it is not clear what is meant by "slope", as it may be different depending on the direction in which we move.

The simplest approach is to keep one variable fixed and study the behaviour of the other one.

Example: $z = f(x, y) = (x-3)^2 + 2xy^2 - 16$.

Set $y = y_0 = 2 \Longrightarrow f(x, y_0) = (x-3)^2 + 8x - 16$. This is now a function of a single variable and, as it is differentiable, we can use differentiation as usual to find its derivative.

For x_0 fixed, we get $f(x_0, y) = (x_0 - 3)^2 + 2x_0y^2 - 16$ and the first derivative w.r.t. y is

$$\frac{\partial}{\partial y} f(x_0, y) = 4x_0 y$$

This illustrates the concept of *partial derivative*. The (first order) partial derivative w.r.t. x of a function of two variables $z = f(x, y)$ is the derivative of the function of one variable $f(x, y_0)$ w.r.t. x (that is, the function z keeping $y = y_0$).

We denote a partial derivative as $\frac{\partial f}{\partial x}$, $\frac{\partial}{\partial x} f(x, y)$, or f_x, and analogously for y.

Example: $f(x, y) = 3x^2y^3 + 4xy + x^2e^{7y}$.
Then $\frac{\partial f}{\partial x} = 6xy^3 + 4y + 2xe^{7y}$ and $\frac{\partial f}{\partial y} = 9x^2y^2 + 4x + 7x^2e^{7y}$.

It goes without saying that we can also consider *higher order partial derivatives*, for example:

$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = f_{xx}$, $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}$,

$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = f_{xy}$, $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y \partial y} = f_{yy}$

Example: $f(x, y) = 2x^4y^3 - x^3y^6$

$f_x = 8x^3y^3 - 3x^2y^6$
$f_y = 6x^4y^2 - 6x^3y^5$
$f_{xy} = 24x^3y^2 - 18x^2y^5$
$f_{yx} = 24x^3y^2 - 18x^2y^5$
$f_{xx} = 24x^2y^3 - 6xy^6$
$f_{yy} = 12x^4y - 30x^3y^4$

Theorem: If $f(x, y)$ is continuous then the second order derivatives f_{xy} and f_{yx} are the same.

Differential, Gradient and Hessian

What is the differential of a function f, df ? Geometrically (in one dimension, see Figure 1): dx times $f'(x_0)$, that is the slope of the tangent at x_0. Then the differential is:

$$df = f'(x_0)\, dx$$

If we have $z = f(x, y)$ we can define two *partial differentials*:

$$df = \frac{\partial f(x_0, y_0)}{\partial x} dx$$
$$df = \frac{\partial f(x_0, y_0)}{\partial y} dy$$

The two df measure changes in f as a response to a change in x (y), if y (x) is held constant.

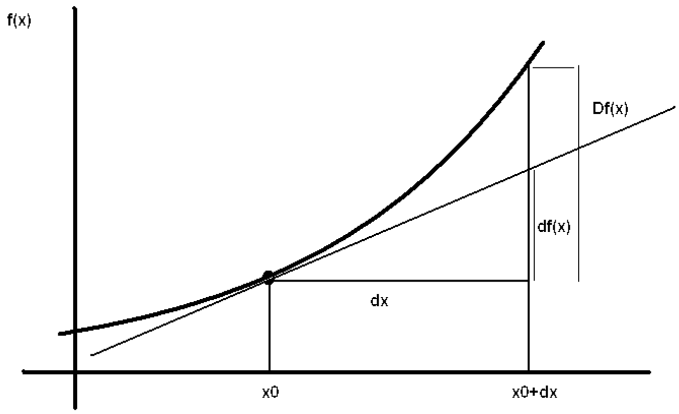

Figure 1: Differential

However, in the multivariate case, what if we change both x and y at the same time, from (x_0, y_0) to $(x_0 + \Delta x, y_0 + \Delta y)$?

We approximate the total change in f $\quad (\Delta f = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0))$ by the sum of one-variable changes:

$$\Delta f \approx df = \frac{\partial f(x_0, y_0)}{\partial x} dx + \frac{\partial f(x_0, y_0)}{\partial y} dy$$

This is called the *total differential*.

The vector $\left(\frac{\partial f(x_0,y_0)}{\partial x}, \frac{\partial f(x_0,y_0)}{\partial y}\right)^T$, whose entries are the first-order partial derivatives, is called the *gradient* at (x_0, y_0). It is often denoted as $Df_{(x_0,y_0)}$.

The two-variable analogue of the second derivative is a symmetric matrix called the *Hessian* matrix, also evaluated at (x_0, y_0). It is often denoted as $D^2 f_{(x_0,y_0)}$. The Hessian is given by:

$$D^2 f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial xy} \\ \frac{\partial^2 f}{\partial yx} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

As it holds that the second partial cross-derivatives are equal, i.e. $\frac{\partial^2 f}{\partial xy} = \frac{\partial^2 f}{\partial yx}$, the Hessian can be written as

$$D^2 f(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

which obviously is a symmetric (2×2) matrix.

Taylor Series Expansion

The so-called Taylor's Theorem provides us with a way of describing a function in terms of its derivatives.

Suppose we know the value of $f(x,y)$ at (x_0, y_0). Suppose also that we can compute up to the 2^{nd} derivative of such function for either argument, obtaining the gradient vector at (x_0, y_0), $Df_{(x_0,y_0)}$, and the Hessian matrix $D^2 f_{(x_0,y_0)}$, also evaluated at (x_0, y_0).

Then Taylor's Theorem (in a simplified form) tells us that:

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) &= f(x_0, y_0) + \\ &+ (\Delta x, \Delta y) \cdot Df(x,y)_{|(x_0,y_0)} + \\ &+ \frac{1}{2} \cdot (\Delta x, \Delta y) \cdot D^2 f(x,y)_{|(x_0,y_0)} \cdot (\Delta x, \Delta y)^T + R \end{aligned}$$

where R is a remainder term.

2 Unconstrained Optimization

Definition (critical points or extrema): the points on the surface of the graph of f where the gradient vector $Df(x,y)$ is the zero vector $\mathbf{0}$ are called critical points of f. Important examples of critical points are local maximum and local minimum points of f.

Note: there are critical points that are neither maxima or minima. We call these points *saddle points*.

How do we assess if a given critical point is a maximum, minimum or a saddle point?

The following "derivative tests" can help identify the nature of a critical point:

1. If $Df(x^*, y^*) = \mathbf{0}$ and $D^2 f(x^*, y^*)$ is a *negative semidefinite* matrix, then the critical point (x^*, y^*) is a *local maximum*;

2. If $Df(x^*, y^*) = \mathbf{0}$ and $D^2 f(x^*, y^*)$ is a *positive semidefinite* matrix, then the critical point (x^*, y^*) is a *local minimum*;

3. Finally, recall from linear algebra that the test for positive/negative semidefiniteness for (2×2) matrices requires the determinant to be non-negative. Hence, if we find that $\det\left(D^2 f(x^*, y^*)\right) <$

~~0 for some~~ critical point (x^*, y^*), then this point must be a saddle point. A saddle point (x^*, y^*) of f is a min of f in some direction, and a max of f in other directions. If, however, we find that $\det\left(D^2 f\left(x^*, y^*\right)\right) = 0$, then further investigation is required.

The rationale of this test can be seen in relation to the Taylor's series expansion of the function we are studying. At a critical point the gradient is the zero vector and, assuming a remainder R close enough to zero, we have that

$$f\left(x_0 + \Delta x, y_0 + \Delta y\right) - f\left(x_0, y_0\right) =$$
$$= \frac{1}{2} \cdot (\Delta x, \Delta y) \cdot D^2 f(x, y)_{|(x_0, y_0)} \cdot (\Delta x, \Delta y)^T$$

It is obvious that if $D^2 f(x, y)_{|(x_0, y_0)}$ is negative semidefinite the quadratic form in the RHS is going to be negative, implying that every movement $(\Delta x, \Delta y)$ away from (x_0, y_0) will decrease the value of f (so we are at a maximum). A similar (but opposite) reasoning can be applied for a minimum.

Example: Find the extrema of $f(x, y) = x^3 - 3x^2 y + 3xy^2 + y^3 - 3x - 21y$.

We compute the gradient vector and set it equal to the zero vector:
$\frac{\partial f}{dx} = 3x^2 - 6xy + 3y^2 - 3$
$\frac{\partial f}{dy} = -3x^2 + 6xy + 3y^2 - 21$

$$\begin{bmatrix} 3x^2 - 6xy + 3y^2 - 3 \\ -3x^2 + 6xy + 3y^2 - 21 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

This gives us a non-linear system of equations. We want to solve for all vectors (x, y) that satisfy it:

$$3x^2 - 6xy + 3y^2 - 3 = 0$$
$$-3x^2 + 6xy + 3y^2 - 21 = 0$$

Adding the two equations together gives

$$0 + 0 + 6y^2 - 24 = 0$$

or

$$y^2 = 4$$

or

$$y = \pm 2$$

So there are two possibilities:

1. $y = 2$: then we solve $3x^2 - 12x + 12 - 3 = 0$ or $x^2 - 4x + 3 = 0 \Longrightarrow x_1 = 3$ and $x_2 = 1$

2. $y = -2$: then we solve $3x^2 + 12x + 9 = 0$ or $x^2 + 4x + 3 = 0 \Longrightarrow x_{1,2} = \frac{-4 \pm \sqrt{16-12}}{2} \Longrightarrow x_1 = -1$ and $x_2 = -3$.

So, ~~the critical points are:~~ $(3, 2)\,,\ (1, 2)\,,\ (-1, -2)\,,\ (-3, -2)\,.$

To check that the first order conditions are sufficient, we also need to get the Hessian matrix:

$f_{xx} = 6x - 6y$
$f_{yy} = 6x + 6y$
$f_{xy} = -6x + 6y$

$$\Longrightarrow D^2 f(x, y) = \begin{bmatrix} (6x - 6y) & (-6x + 6y) \\ (-6x + 6y) & (+6x + 6y) \end{bmatrix}$$

Now we have to check the Hessian matrix at all critical points:

1. Point $(3, 2)$: $\begin{bmatrix} 6 & -6 \\ -6 & 30 \end{bmatrix}$ and $6 > 0$, $30 > 0$, $180 - 36 > 0 \Longrightarrow D^2 f(3, 2)$ is positive definite, so this is a minimum.

2. Point $(1, 2)$: $\begin{bmatrix} -6 & 6 \\ 6 & 18 \end{bmatrix}$ and we can check that this matrix is neither positive nor negative semidefinite, so this is a saddle point.

3. Point $(-3, -2)$: $\begin{bmatrix} -6 & 6 \\ 6 & 30 \end{bmatrix}$ $D^2 f(-3, -2)$ is negative definite, so this is a maximum.

4. Point $(-1, -2)$: $\begin{bmatrix} 6 & -6 \\ -6 & -18 \end{bmatrix}$ this matrix is neither positive nor negative semidefinite, so this is a saddle point.

See math textbooks for the description of methods using Gradient/Hessian to determine if a function $f(x, y)$ is concave/convex.

Exercise: Find the extrema of $f(x, y) = 3x^2 + 3xy + 3y^2 - 9x + 1$.

Answer: Local Minimum at $(x, y) = (2, -1)$, as $D^2 f(2, -1)$ is positive definite.

3 Principles of Constrained Optimization with (Linear) Equality Constraints

The general structure of a constrained optimization problem is represented as follows:

$$\max_{(x,y)} f(x,y) \quad s.t.\ g(x,y) = 0$$

e.g., the maximization of a utility function subject to the budget constraint.[1]

We now introduce the concept of the *Lagrangean* function and define it as:[2]

$$L(x,y,\lambda) = f(x,y) - \lambda g(x,y)$$

A natural question to ask is the following: what is the relationship between the constrained optimization problem and the Lagrangean function? The answer is: if (x^*, y^*) solves the constrained optimization problem

$$\max_{(x,y)} f(x,y) \quad s.t.\ g(x,y) = 0$$

then there exists some real number λ^* such that $L(x,y,\lambda) = f(x,y) - \lambda g(x,y)$ has a critical point at (x^*, y^*, λ^*) .

The existence of such critical point at (x^*, y^*, λ^*) is **NECESSARY** for (x^*, y^*) to be a constrained max of the original problem, but it is **not sufficient** (please see the discussion of why this is the case in a math textbook).

Example 1: $\max_{(x,y)} 4xy - 2x^2 + y^2$ subject to $g(x,y) = 3x + y - 5 = 0$.

Write the Lagrangean:

$$L(x,y,\lambda) = 4xy - 2x^2 + y^2 - \lambda(3x + y - 5)$$

[1] Notice that we focus on the max problem. This is without loss of generality as we can always express a min problem as the maximization of minus the original function. Namely, $\arg\min_{(x,y)} f(x,y) = \arg\max_{(x,y)} [-f(x,y)]$.

[2] We can also define the Lagrangean as $L\left(x,y,\widehat{\lambda}\right) = f(x,y) + \widehat{\lambda} g(x,y)$. In this formulation, the multipliers $\widehat{\lambda}$ are opposite in sign compared to the alternative formulation with the multipliers λ and a minus sign in front of the term $\lambda g(x,y)$.

Compute the gradient vector:

$L_x = 4y - 4x - 3\lambda$

$L_y = 4x + 2y - \lambda$

$L_\lambda = -(3x + y - 5)$

Solve the system:

$$4y - 4x - 3\lambda = 0$$
$$4x + 2y - \lambda = 0$$
$$-(3x + y - 5) = 0$$

The solution, for example using Gauss-Jordan, is: $x^* = -1,\ y^* = 8,\ \lambda^* = 12$.

The issue now is: how can we check if the point $(x^*, y^*) = (-1, 8)$ corresponds to a constrained max of f? This question is especially relevant when we have many critical points of L.

In order to assess whether a specific critical point of L is also a constrained optimum of f we can use a method that, although very tedious, provides conditions for the case of 2 (or more) variables and 1 (or more) constraints.[3]

The method is a typical Second Order Conditions method. In particular, we expect the second order condition to involve the negative definiteness of a quadratic form along a linear constraint set.

First, we form the so-called *bordered Hessian* H, i.e. $D^2L(x, y, \lambda)$ that is defined as follows:

$$H = \begin{bmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{xy} & L_{yy} \end{bmatrix}$$

H is nothing but the second derivative of L w.r.t. x and y, with the added "borders" given by the derivative of the constraints with respect to the same x and y.

If we had more than one constraint (say k constraints) and we were maximising over a set of n

[3] Here we describe such method in the (2×1) environment for the example above. See, among others, Simon and Blume "Mathematics for Economists" for more details and several other examples.

variables (x_1 to x_n), the bordered Hessian might in general look like:

$$H = \begin{bmatrix} 0 & \dots & 0 & g_{1x_1} & \cdots & g_{1x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & g_{kx_1} & \cdots & g_{kx_n} \\ g_{1x_1} & \cdots & g_{kx_1} & L_{x_1x_1} & \cdots & L_{x_1x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_{1x_n} & \cdots & g_{kx_n} & L_{x_1x_n} & \cdots & L_{x_nx_n} \end{bmatrix}$$

Which submatrices of H do we have to consider?

As a practical direction, keep in mind that we want to check the signs of the *last* $(\mathbf{n}-\mathbf{k})$ leading principal minors of H, starting with the determinant of H itself. Different cases are possible:

1. if $\det H$ has the same sign as $(-1)^n$ and if the last $n-k$ leading principal minors alternate in sign, then the matrix of L's second derivatives

$$\begin{bmatrix} L_{x_1x_1} & \cdots & L_{x_1x_n} \\ \vdots & \ddots & \vdots \\ L_{x_1x_n} & \cdots & L_{x_nx_n} \end{bmatrix}$$

is Negative definite and the critical point we are considering is a Max over the constraint set.

2. if both $\det H$ and the last $n-k$ leading principal minors have all the same signs $(-1)^k$ then the matrix of L's second derivatives is positive definite and the critical point is a Min on the constraint set;

3. if both the previous conditions are violated by *non-zero* leading principal minors, then the matrix of L's second derivatives is indefinite and the critical point is neither a Max nor a Min on the constraint set.[4]

Going back to the initial example: $L_{xx} = -4,\ L_{yy} = +2,\ L_{xy} = 4,\ g_x = 3,\ g_y = 1$, so that

$$H = \begin{bmatrix} 0 & 3 & 1 \\ 3 & -4 & 4 \\ 1 & 4 & 2 \end{bmatrix}$$

[4]For more details on this topic, please refer to the book by Simon and Blume, from page 388 onwards.

In the simple two variables example, if the bordered Hessian has a determinant with the same sign as $(-1)^n$, where $n\,(=2$ in the example) is the number of exogenous variables in the original optimization problem, then $(x^*, y^*)\,(=(-1,8)$ in the example) is the constrained maximum of f. In the example:

$$\det H = 10,\ (-1)^{n=2} > 0$$

so we know that this critical point corresponds to a max.

Example 2: $f(x,y) = 2xy,\ g(x,y) = 3x + 2y - 60 = 0$

$$L = 2xy - \lambda(3x + 2y - 60)$$

$L_x = 2y - 3\lambda = 0 \Longrightarrow y = \frac{3}{2}\lambda$
$L_y = 2x - 2\lambda = 0 \Longrightarrow x = \lambda$
$L_\lambda = -(3x + 2y - 60) = 0 \Longrightarrow \lambda = x = \frac{2}{3}y$

$\Longrightarrow 3\frac{2}{3}y + 2y - 60 = 0 \Longrightarrow 4y = 60 \Longrightarrow y = 15 \Longrightarrow x = 10$

$$H = \begin{bmatrix} 0 & 3 & 2 \\ 3 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$$

and $\det H = 24 > 0$, so that we know it is a Max.

Example 3: $\max_{(x,y,z)} x^2 + 3y^2 + 2z^2$ subject to $x + 3y = 30$ and $y + 2z = 20$

$$L = x^2 + 3y^2 + 2z^2 - \lambda_1(x + 3y - 30) - \lambda_2(y + 2z - 20)$$

$L_x = 2x - \lambda_1 = 0 \Longrightarrow x = \frac{1}{2}\lambda_1$
$L_y = 6y - 3\lambda_1 - \lambda_2 = 0$
$L_z = 4z - 2\lambda_2 = 0 \Longrightarrow \lambda_2 = 2z$

$\Longrightarrow 6y - 6x - 2z = 0$

$L_{\lambda_1} = x + 3y - 30 = 0$
$L_{\lambda_2} = y + 2z - 20 = 0 \Longrightarrow y = 20 - 2z$

then $6y - 6x - 2z = 0$ and $y = 20 - 2z \Longrightarrow 3(20 - 2z) - 3x - z = 0$

$\Longrightarrow 60 - 6z - 3x - z = 0$ and

$$60 = 7z + 3x$$

Also, $x + 3y - 30 = 0$ and $y = 20 - 2z \Longrightarrow x + 3(20 - 2z) - 30 = 0$

$\Longrightarrow x + 60 - 6z - 30 = 0$ and

$$x = 6z - 30$$

Finally, $x = 6z - 30$ and $60 = 7z + 3x \Longrightarrow\ 60 = 7z + 18z - 90$

$\Longrightarrow 150 = 25z \Longrightarrow\ z^* = 6 \Longrightarrow x^* = 6$ and $\ y^* = 8$. Also, $\lambda_1^* = 12$ and $\lambda_2^* = 12$.

We can now look at the SOC's:

$L_{xx} = 2$
$L_{xy} = 0$
$L_{xz} = 0$
$L_{yy} = 6$
$L_{yx} = 0$
$L_{yz} = 0$
$L_{zz} = 4$
$L_{zy} = 0$
$L_{zx} = 0$

$\frac{\partial g_1}{\partial x} = 1,\ \frac{\partial g_1}{\partial y} = 3,\ \frac{\partial g_1}{\partial z} = 0$

$\frac{\partial g_2}{\partial x} = 0,\ \frac{\partial g_2}{\partial y} = -1,\ \frac{\partial g_2}{\partial z} = 2$

So the bordered Hessian is

$$H = \begin{bmatrix} 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 1 & 0 & 2 & 0 & 0 \\ 3 & 1 & 0 & 6 & 0 \\ 0 & 2 & 0 & 0 & 4 \end{bmatrix}$$

Here we have 3 unkowns and 2 constraints: $n - k = 3 - 2 = 1$. $(-1)^{n=3} = -1 \implies$ so, check as an exercise if $\det H < 0$ to see if it is a max (or something else: Min? Saddle point?)

4 Non Linear Programming

Non linear programming allows us to deal with constrained optimization problems where both the objective function $f(x,y)$ and the inequality constraints are *non-linear*.

An example of such a problem is:[5]

$$\max_{(x,y)} 2x + y = z$$

subject to

$$\begin{aligned} &-x^2 + 4x - y \leq 0 \\ &2x + 3y \leq 12 \\ &x, y \geq 0 \end{aligned}$$

To make things easier in terms of notation, we will restrict attention to optimization problems with non-negativity constraints only. The method introduced here is the same as for problems with non-linear inequality constraints, and is also readily generalized to the case of n exogenous variables $x_1, x_2, ..., x_n$ and k constraints $g_1(x_1, ..., x_n), ..., g_k(x_1, ..., x_n)$.

The type of problem we will study is then:

$$\max_{(x,y)} f(x,y)$$

subject to

$$\begin{aligned} x &\geq 0 \\ y &\geq 0 \end{aligned}$$

In the above setup we have:[6]

$$\begin{aligned} g_1(x,y) &= -x, \;\; c_1 = 0 \\ g_2(x,y) &= -y, \;\; c_2 = 0 \end{aligned}$$

which means that the constraints are rewritten as:

$$\begin{aligned} -x &\leq 0 \\ -y &\leq 0 \end{aligned}$$

[5] In this simple case the graph of the objective function is a line (an isoquant, with changing intercept z). In fact, $2x + y = z \Longrightarrow y = z - 2x$.

[6] It is convention to denote constraints as $g_i(x_1, ..., x_n) \leq c_i$, $i = 1, ..., k$. Alternatively, we can denote the constraints as $g_i(x_1, ..., x_n) \geq c_i$, $i = 1, ..., k$, remembering that this is going to change the sign of the multipliers compared to the other formulation, if in the Lagrangean there is a minus sign in front of the term $\lambda g(x,y)$.

We can set up a Lagrangean function, as we did in the case of equality constraints. In this general case it will read as follows:

$$L(x, y, \lambda_1, ..., \lambda_k) = f(x, y) - \lambda_1 [g_1(x, y) - c_1] - ... - \lambda_k [g_k(x, y) - c_k]$$

And in our particular problem, with non-negativity constraints only, the Lagrangean becomes:

$$L(x, y, \lambda_1, \lambda_2) = f(x, y) - \lambda_1 [-x] - \lambda_2 [-y] = f(x, y) + \lambda_1 x + \lambda_2 y.$$

We will now provide the so-called Kuhn-Tucker Conditions (KTC) for problems of this type. These are a mix of fairly standard FOC's coupled with *complementary slackness conditions*.

$\frac{\partial L}{\partial x} \leq 0,\ x \cdot \frac{\partial L}{\partial x} = 0$

$\frac{\partial L}{\partial y} \leq 0,\ y \cdot \frac{\partial L}{\partial y} = 0$

$\frac{\partial L}{\partial \lambda_1} \geq 0,\ \lambda_1 \cdot \frac{\partial L}{\partial \lambda_1} = 0$

$\frac{\partial L}{\partial \lambda_2} \geq 0,\ \lambda_2 \cdot \frac{\partial L}{\partial \lambda_2} = 0$

Example 1:

$$\max_{(x,y)} 1 - 8x + 10y - 2x^2 - 3y^2 + 4xy$$

subject to

$$\begin{aligned} x &\geq 0 \\ y &\geq 0 \end{aligned}$$

We start by writing the Lagrangean as:

$$L(x, y, \lambda_1, \lambda_2) = 1 - 8x + 10y - 2x^2 - 3y^2 + 4xy + \lambda_1 x + \lambda_2 y$$

Then, the KTC's are:

$$
\begin{aligned}
\frac{\partial L}{\partial x} &= -8-4x+4y+\lambda_1 \le 0 && (1)\\
\frac{\partial L}{\partial y} &= 10-6y+4x+\lambda_2 \le 0 && (2)\\
\frac{\partial L}{\partial \lambda_1} &= x \ge 0 && (3)\\
\frac{\partial L}{\partial \lambda_2} &= y \ge 0 && (4)\\
\frac{\partial L}{\partial x}\cdot x &= -8x-4x^2+4xy+\lambda_1 x = 0 && (5)\\
\frac{\partial L}{\partial y}\cdot y &= 10y-6y^2+4xy+\lambda_2 y = 0 && (6)\\
\frac{\partial L}{\partial \lambda_1}\cdot \lambda_1 &= \lambda_1 x = 0 && (7)\\
\frac{\partial L}{\partial \lambda_2}\cdot \lambda_2 &= \lambda_2 y = 0 && (8)
\end{aligned}
$$

Notice that at an optimum all conditions (1)-(8) have to be satisfied simultaneously, and that conditions (7) and (8) are satisfied only in the following cases.

Case 1: Suppose $\lambda_1 = \lambda_2 = 0$. The KTC set of restrictions reduces to:

Condition $(5):\ -8x-4x^2+4xy=0 \Longrightarrow\ -8-4x+4y=0$
Condition $(6):\ 10y-6y^2+4xy=0\ \Longrightarrow 10-6y+4x=0$

The conditions above can be respectively rewritten as $x = y-2$ and $3y = 5+2x$.

Then combining the two gives $5-3y+2(y-2) = 1-y = 0 \Longrightarrow\ y^* = 1$.

This however would imply $x^* = 1-2 = -1$ which is a violation of the non negativity constraint. Since we ended up with a contradiction, the assumption $\lambda_1 = \lambda_2 = 0$ is wrong, and we can rule out the possibility that the pair $(x^* = -1, y^* = 1)$ is an optimum.

Case 2: Suppose $\lambda_1 \ge 0,\ \lambda_2 = 0$. This implies that $x^* = 0$. Then the KTC set of restrictions reduces to:

Condition $(1):\ -8+4y+\lambda_1 \le 0$

Condition $(2):\ 10-6y \le 0$ or $y \ge \frac{10}{6} = \frac{5}{3}$

Condition $(5):\ 0 = 0$

Condition (6) : $10y - 6y^2 = 0$ or $10 - 6y = 0$ or $y^* = \frac{5}{3}$

So we have $x^* = 0,\ y^* = \frac{5}{3},\ \lambda_2^* = 0$.

What about λ_1^*?

$-8 + 4\frac{5}{3} + \lambda_1 \leq 0 \Longrightarrow \lambda_1 \leq 8 - \frac{20}{3} \Longrightarrow \lambda_1 \leq \frac{24-20}{3} = \frac{4}{3}$.

So we conclude that $\lambda_1 \in \left(0, \frac{4}{3}\right]$.

Case 3: Suppose $\lambda_1 = 0,\ \lambda_2 \geq 0$. This implies that $y^* = 0$. Then the KTC set of restrictions reduces to:

Condition (1) : $-8 - 4x \leq 0$ or $x \leq -2$ so we have another violation of the constraint.

So we exclude the possibility that $\lambda_2 \geq 0$ and we stop here.

Case 4: Suppose $\lambda_1 \geq 0,\ \lambda_2 \geq 0$. This implies that $x^* = y^* = 0$. However, from condition (2) we get that $10 + \lambda_2 \leq 0$ or $\lambda_2 \leq -10$, which is a contradiction and we stop here.

By exclusion, the solution is case 2.

Example 2:

$$\max_{(x,y)} 2\sqrt{x} + 2\sqrt{y}$$

subject to

$$\begin{aligned} 2x + y &\leq 3 \\ x + 2y &\leq 3 \end{aligned}$$

The graphical representation of the feasible set is rather straightforward (see Figure 2). It goes without saying that we are dealing with a real-valued function, hence its domain is restricted to $x \geq 0$ and $y \geq 0$. However, there is no need to include these additional inequality constraints in the maximization problem. If a candidate solution were to violate them, it would be discarded.

We can then write the Lagrangean as

$$L(x, y, \lambda_1, \lambda_2) = 2\sqrt{x} + 2\sqrt{y} - \lambda_1 (2x + y - 3) - \lambda_2 (x + 2y - 3)$$

A lot of time is saved by looking carefully if any of the constraints is binding.

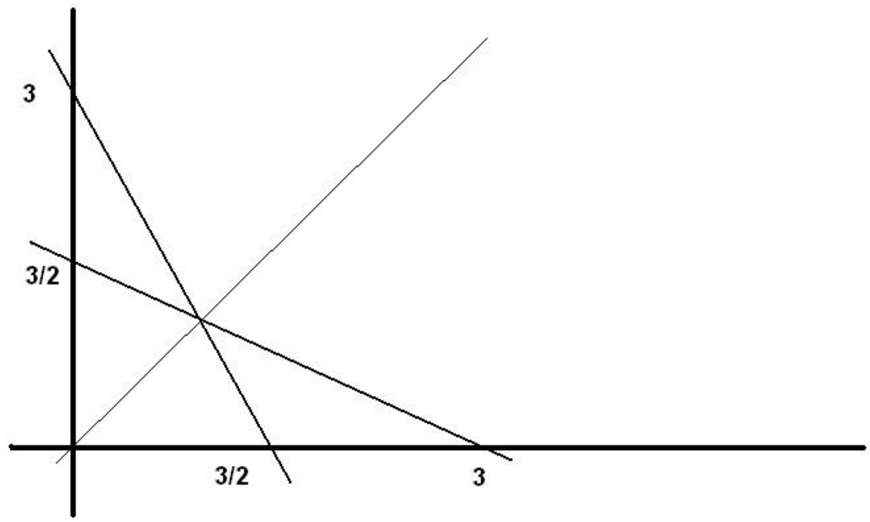

Figure 2: Graph of the symmetric constraints.

1. If both constraints are NOT binding at the optimum, no max would exist, as $f(x, y)$ is strictly (monotonically) increasing in both x and y. $\left(\frac{\partial f}{\partial x} = \frac{1}{\sqrt{x}} > 0 \; \forall x\right)$.

2. If only one constraint were binding (e.g., $2x + y = 3$) then we could reduce the problem to:

$$\max_x 2\sqrt{x} + 2\sqrt{3 - 2x}$$

and the FOC yields: $x = \frac{1}{2}$, $y = 2$. This would however violate the second constraint $x + 2y = \frac{1}{2} + 4 > 3$. So this is not possible.

3. Hence, both constraints must be binding, as it is evident from the graph.