4: Optimization

Math camp 2019

George Mason University

August 22, 2019

- Me: Arthur Dolgopolov,
- Don't hesitate to email me questions! Seriously.
  adolgopo@gmu.edu
- Lecture slides:
  https://arthurdolgopolov.net/teaching/mathcamp
- Wolfram Mathematica (free for students):
  https://cos.gmu.edu/mathematica/

- 0. Finish linear algebra
- 1. Unconstrained optimization
- 2. Optimization with constraints (Lagrange)

- Math for CS (chapters 15,17) https://courses.csail.mit.edu/6.042/spring17/mcs.pdf
- Mathematical methods for economic theory Martin J. Osborne (chapter 7, but also 2,3,4,5,6) https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/toc

Taylor series

We have built intuition for the derivative as an approximation at that point. More generally, the more the derivatives of the functions coincide the more one approximates the other in the neighborhood. Formally, this is Taylor series.

Taylor series for infinitely differentiable $f(x)$:

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + ..,$$

When $a = 0$ this is called Maclaurin series.

Practice:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} ...$$

$$sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} ...$$

General chain rule and total derivative

If g^j is differentiable function of m variables and f is differentiable function of n variables, and the function F of m variables is defined by

$F(x_1, ..., x_m) = f(g^1(x_1, ..., x_m), ..., g^n(x_1, ..., x_m))$ for all $(x_1, ..., x_m)$

then F is differentiable and for $k = 1, ..., m$ we have

$F_k(x_1, ..., x_m) = \sum_{i=1}^{n} f_i(g^1(x_1, ..., x_m), ..., g^n(x_1, ..., x_m))g_k^i(x_1, ..., x_m)$,

where g_k^i is the partial derivative of g^i with respect to its k-th argument.

Implicit differentiation:

x is defined implicitly
$f(x, p) = 0$, or $f(x(p), p) = 0$.
Differentiate using chain rule:

$$f_1'(x(p), p)x'(p) + f_2'(x(p), p) = 0$$

$$x'(p) = -f_2'(x(p), p) / f_1'(x(p), p)$$

if $f_1'(x(p), p) \neq 0$.

Practice:

The function g is defined implicitly by the condition $F(f(x,y), g(y)) = h(y)$. Find the derivative $g'(y)$ in terms of the functions F, f, g, and h and their derivatives.

Theorem (Implicit Function theorem)

Let f be a continuously differentiable function of two variables defined on an open set S. If $f_2'(x_0, y_0) \neq 0$ then there exists a continuously differentiable function g of a single variable defined on an open interval I containing x_0 such that $f(x, g(x)) = f(x_0, y_0)$ for all $x \in I$, and $g'(x_0)$, the slope of the level curve of f for the value $f(x_0, y_0)$ at the point (x_0, y_0), is

$$-\frac{f_1'(x_0, y_0)}{f_2'(x_0, y_0)}$$

Application: level curves
https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/dfi/t

Consider again an economic example:

$$Y = C + I + G$$

$$C = f(Y - T)$$

$$I = h(r)$$

$$r = m(M)$$

The variables are Y (national income), C (consumption), I (investment) and r (the rate of interest), and the parameters are M (the money supply), T (the tax burden), and G (government spending).
Find a market equilibrium in a system of equations.

Step 1: Totally differentiate the system
Step 2: Solve using Cramer's rule / Gaussian elimination / Manually
https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/dif/t

the Jacobian matrix, the total derivative, the gradient.
$J = [\frac{\partial f}{\partial x_1} ... \frac{\partial f}{\partial x_n}] = (\nabla f)^T$
Widely applied in economics to study dynamics or equilibrium of systems, e.g. systems of linear equations.

A function is homogeneous of degree k

$f(tx_1, ...tx_n) = t^k f(x_1, ...x_n)$ for all $(x1, ...x_n) \in S$ and $t > 0$.

What if $k = 0$?

Part 1: Unconstrained optimization

Common problem in microeconomics, **consumer problem**:

$$max_a u(a) \text{ s.t. } a \in S$$

or maybe, **production problem**:

$$min_a c(a) \text{ s.t. } a \in S$$

or even:

$$max \ u_1(\mu) + u_2(\mu) \text{ s.t. } \mu \in S$$

In mathematics these are **optimization problems**

The optimization problems we study take the form

$$max\ f(x) \text{ s.t. } x \in S$$

where f is a function, x is an n-vector (which we can also write as $(x_1, ..., x_n)$), and S is a set of n-vectors.

- f - objective function,
- x - the choice variable,
- S - the constraint set or opportunity set or feasible set.

In contrast with feasibility problems

$$find\ x \in U \text{ s.t. } x \in S$$

Let f be a function of many variables defined on a set X and let S be a subset of X.
The point $x^* \in S$ solves the problem

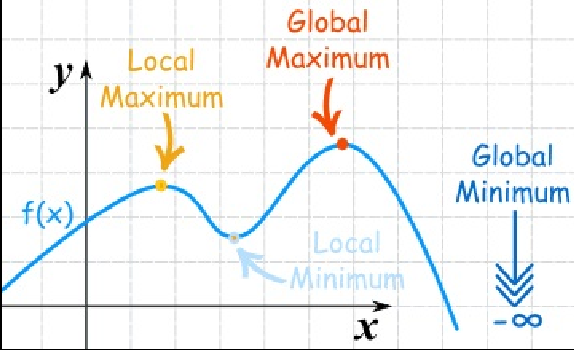
$$\max f(x) \ s.t. \ x \in S$$

if

$$f(x) \leq f(x^*) \forall x \in S.$$

In this case we say that x^* is a **maximizer** of $f(x)$ s.t. the constraint $x \in S$, and that $f(x^*)$ is the **maximum** (or maximum value) of $f(x)$ s.t. the constraint $x \in S$.

An optimization problem may have many maximizers and minimizers, but has at most one maximum and at most one minimum.


Global
Maximum
Local
Maximum
y
f(x)
Local
Minimum
x
Global
Minimum
-∞

Let f be a function of many variables defined on a set X and let S be a subset of X.

Definition

The point $x^* \in S$ is a local maximizer of $f(x)$ subject to the constraint $x \in S$ if there is a number $\epsilon > 0$ such that $f(x) \leq f(x^*)$ for all $x \in S$ for which the distance between x and x^* is at most ϵ.

Definition

A set is compact if it is closed and bounded.

Definition

A set is bounded if there exists a number k such that the distance of every point in S from the origin is at most k.

Definition

The set S of n-vectors is open if every point in S is an interior point of S. The set S of n-vectors is closed if every boundary point of S is a member of S.

Definition

A point x is a boundary point of a set S of vectors if for every number $\epsilon > 0$ (however small), at least one point within the distance ϵ of x is in S, and at least one point within the distance ϵ of x is outside S.

Theorem

Let f be a function of many variables defined on a set X and let S be a subset of X. If f is continuous and S is compact then the problems of maximizing and minimizing $f(x)$ subject to $x \in S$ both have solutions.

Theorem

Let f be a function of many variables defined on a set X and let S be a subset of X. If f is continuous and S is compact then the problems of maximizing and minimizing $f(x)$ subject to $x \in S$ both have solutions.

Note! Only sufficiency.

- $f(x) = x, S = [0, \infty)$ (i.e. S is the set of all nonnegative real numbers). In this case, f increases without bound, and never attains a maximum.

- $f(x) = x, S = [0, \infty)$ (i.e. S is the set of all nonnegative real numbers). In this case, f increases without bound, and never attains a maximum.
- $f(x) = 1 - 1/x, S = [1, \infty)$. In this case, f converges to the value 1, but never attains this value.

- $f(x) = x, S = [0, \infty)$ (i.e. S is the set of all nonnegative real numbers). In this case, f increases without bound, and never attains a maximum.
- $f(x) = 1 - 1/x, S = [1, \infty)$. In this case, f converges to the value 1, but never attains this value.
- $f(x) = x, S = (0, 1)$. In this case, the points 0 and 1 are excluded from S (which is an open interval). As x approaches 1, the value of the function approaches 1, but this value is never attained for values of x in S, because S excludes x = 1.

- $f(x) = x, S = [0, \infty)$ (i.e. S is the set of all nonnegative real numbers). In this case, f increases without bound, and never attains a maximum.
- $f(x) = 1 - 1/x, S = [1, \infty)$. In this case, f converges to the value 1, but never attains this value.
- $f(x) = x, S = (0, 1)$. In this case, the points 0 and 1 are excluded from S (which is an open interval). As x approaches 1, the value of the function approaches 1, but this value is never attained for values of x in S, because S excludes x = 1.
- $f(x) = x$ if $x < 1/2$ and $f(x) = x - 1$ if $x \geq 1/2; S = [0, 1]$. In this case, as x approaches $1/2$ the value of the function approaches $1/2$, but this value is never attained, because at $x = 1/2$ the function jumps down to $-1/2$.

The usual approach to solving the optimization problem is to take a derivative.

Definition

Let f be a function of a single variable defined on a set S. A point $x \in S$ at which f is differentiable and $f'(x) = 0$ is a stationary point of f.

- Is it necessary/sufficient for the solution to optimization problem to be a stationary point?

The usual approach to solving the optimization problem is to take a derivative.

Definition

Let f be a function of a single variable defined on a set S. A point $x \in S$ at which f is differentiable and $f'(x) = 0$ is a stationary point of f.

- Is it necessary/sufficient for the solution to optimization problem to be a stationary point?
- In other words, can we have $f'(x^*) = 0$ with $f(x^*) < f(x)$ for some x?

The usual approach to solving the optimization problem is to take a derivative.

Definition

Let f be a function of a single variable defined on a set S. A point $x \in S$ at which f is differentiable and $f'(x) = 0$ is a stationary point of f.

- Is it necessary/sufficient for the solution to optimization problem to be a stationary point?
- In other words, can we have $f'(x^*) = 0$ with $f(x^*) < f(x)$ for some x?
- Can we have $f'(x^*) \neq 0$ with $f(x^*) \geq (x)$ for all x?

Theorem

Let f be a function of a single variable defined on the interval I. If a point x in the interior of I is a local maximizer or minimizer of f and f is differentiable at x then $f'(x) = 0$.

Theorem

Let f be a function of a single variable defined on the interval I. If a point x in the interior of I is a local maximizer or minimizer of f and f is differentiable at x then $f'(x) = 0$.

Usually, we take derivative to find minimum and maximum of a function. This is called First Order Condition (FOC).
$f'(x^*) = 0$ or for partials $f'_{x_i}(x^*) = 0$ for all x_i, $\nabla f(x) = 0$.
This will not always work because:
1. x^* could be only a local maximum or a local minimum, not a global maximum/minimum.
2. It could be a stationary point that is neither a local maximum nor a minimum (inflection).

$$max\ f(x, y) = x^3 + y^3$$

3. The solution could be on the boundary (corner solution).

Theorem (First Order Conditions)

Let f be a function of n variables defined on the set S. If the point x in the interior of S is a local maximizer or minimizer of f and $f_i'(\cdot)$ exists at x then

$$f_i'(x) = 0$$

The procedure is:

- 1. Inspect the function and show that solution is interior. Or restrict yourself to interior solutions and ignore the boundary.

$$min\ (x-2)^2+(y-1)^2$$

The procedure is:

- 1. Inspect the function and show that solution is interior. Or restrict yourself to interior solutions and ignore the boundary.

$$min\ (x-2)^2 + (y-1)^2$$

- 2. Check the second derivative at x^*. For multivariate functions, this is a lengthy procedure of checking the Hessian. This will tell you if the point is a local maximum or a minimum. It can also be inconclusive.

The procedure is:

- 1. Inspect the function and show that solution is interior. Or restrict yourself to interior solutions and ignore the boundary.

$$min\ (x-2)^2 + (y-1)^2$$

- 2. Check the second derivative at x^*. For multivariate functions, this is a lengthy procedure of checking the Hessian. This will tell you if the point is a local maximum or a minimum. It can also be inconclusive.
- 3. Check the second derivative (Hessian) everywhere. If it works, you may find that f is concave or convex (everywhere). Then local maximum/minimum that you found at previous step is a (global) maximum or minimum.

Definition (Convexity)

f is called **convex** if

$$\forall x_1, x_2 \in X, \forall t \in [0,1] : f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

Definition (Quasiconvexity)

f is called **quasiconvex** if

$$\forall x_1, x_2 \in X, \forall t \in [0,1] : f(tx_1 + (1-t)x_2) \leq max(f(x_1), f(x_2))$$

$$z = -(x^2 + y^2)^{0.2}$$

$$z = -(x^2 + y^2)$$

Visualizing helps:

https://www.google.com/search?q=z+%3D+-%28x%5E2+%2B+y%5E2%29%5E0.5&ie=utf-8&oe=utf-8&client=firefox-b-1

We can summarize the last point in a theorem:

Theorem (Global optimization)

Let f be a differentiable function of a single variable defined on the interval I, and let x be in the **interior** *of I. Then*

- *if f is concave then x is a global maximizer of f in I if and only if x is a stationary point of f (i.e. $f'(x) = 0$)*
- *if f is convex then x is a global minimizer of f in I if and only if x is a stationary point of f (i.e. $f'(x) = 0$).*

The procedure:

Necessary condition if not on the boundary, first thing to check:
$f_i' = 0 \ \forall x_i$ (FOC)

Check if a point is a minimum/maximum (SOC):

$$\begin{bmatrix} x'Ax < 0 \forall x \neq 0 \ \textit{or} \\ \text{Sylvester' criterion } (-1)^k det(A_k) > 0 \text{ for } k = 1, 2..n \text{ or} \\ \text{all eigenvalues are negative} \end{bmatrix} \iff$$

$H(x^*)$ is negative definite $\implies$ local maximum

And similarly for local minimum.

Theorem (Second Order Conditions)

Let f be a twice-differentiable function of n variables with continuous partial derivatives and cross partial derivatives, defined on the set S. Suppose that $f_i'(x^*) = 0$ for $i = 1, ..., n$ for some x^* in the interior of S (so that x^* is a stationary point of f). Let H be the Hessian of f.

- *If $H(x^*)$ is negative definite then x^* is a local maximizer of f.*
- *If $H(x^*)$ is positive definite then x^* is a local minimizer of f.*
- *If x^* is a local maximizer of f then $H(x^*)$ is negative semidefinite.*
- *If x^* is a local minimizer of f then $H(x^*)$ is positive semidefinite.*

$Q = x^T Ax$ is called a quadratic form

Definition

Let Q be a quadratic form, and let A be the symmetric matrix that represents it (i.e. $Q(x) = x'Ax$ for all x).
Q and A are called:

- **positive definite** if $x'Ax > 0$ for all $x \neq 0$
- **negative definite** if $x'Ax < 0$ for all $x \neq 0$
- **positive semidefinite** if $x'Ax \geq 0$ for all x
- **negative semidefinite** if $x'Ax \leq 0$ for all x
- **indefinite** if it is neither positive nor negative semidefinite (i.e. if $x'Ax > 0$ for some x and $x'Ax < 0$ for some x).

- a scalar $f'(x)$ turned into a vector ∇f

- a scalar $f'(x)$ turned into a vector ∇f
- a scalar $f''(x)$ turns into a matrix $\mathbf{H}$, called the "Hessian".

- a scalar $f'(x)$ turned into a vector ∇f
- a scalar $f''(x)$ turns into a matrix $\mathbf{H}$, called the "Hessian".
- $\mathbf{H} = \begin{bmatrix} f''_{11}(x) & f''_{12}(x) & f''_{13}(x) \\ f''_{21}(x) & f''_{22}(x) & f''_{23}(x) \\ f''_{31}(x) & f''_{32}(x) & f''_{33}(x) \end{bmatrix}$

Three ways to get to definiteness:
First way is to look at quadratic form $x^T Ax$ directly.
A quadratic form is a way to write down a polynomial of power 2.
Instead of doing a multiplication, treat dimensions as variables.
Simplest case is diagonal matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

But diagonal elements are just squares $x^2 + y^2$, so the matrix is positive definite.

$$\begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$$

This is just $Q = x^2 + 4y^2 - 2xy = (x - y)^2 + 3y^2$, which is always nonnegative. For all nonzero vectors it is strictly positive. So the matrix is positive definite.

2nd way is through eigenvalues:

Theorem

A matrix is positive definite $\iff$ *all its eigenvalues are positive*
A matrix is negative definite $\iff$ *all its eigenvalues are negative*

As you recall there is a relationship between determinant and eigenvalues. So you already know what the determinant of a definite matrix should be.

3rd way is Sylvester' criterion:

Theorem

Let A be an $n \times n$ symmetric matrix, and let A_k be the submatrix of A taking the upper-left corner of $k \times k$ elements. Denote $det(A_k)$, the leading principal minor. Then:

- *A is positive definite iff $det(A_k) > 0$ for $k = 1, 2...n$*
- *A is negative definite iff $(-1)^k det(A_k) > 0$ for $k = 1, 2...n$*
- *(+) indefinite if and only if $|A| < 0$*

Simpler examples:

- $$f(x, y) = 3xy - x^2 - y^2$$

Simpler examples:

- $$f(x, y) = 3xy - x^2 - y^2$$
- one s. point (0,0)

Simpler examples:

- $$f(x, y) = 3xy - x^2 - y^2$$
- one s. point (0,0)
- $$f(x, y) = x^3 + y^3 - 3xy$$

Simpler examples:

- $$f(x,y) = 3xy - x^2 - y^2$$
- one s. point (0,0)
- $$f(x,y) = x^3 + y^3 - 3xy$$
- (0,0) and (1,1)

Simpler examples:

- $$f(x, y) = 3xy - x^2 - y^2$$
- one s. point (0,0)
- $$f(x, y) = x^3 + y^3 - 3xy$$
- (0,0) and (1,1)
- $$f(x, y) = 8x^3 + 2xy - 3x^2 + y^2 + 1$$

Simpler examples:

- $$f(x, y) = 3xy - x^2 - y^2$$
- one s. point (0,0)
- $$f(x, y) = x^3 + y^3 - 3xy$$
- (0,0) and (1,1)
- $$f(x, y) = 8x^3 + 2xy - 3x^2 + y^2 + 1$$
- (0,0),(1/3,-1/3)

Example:
The volcano

$$x^2 + y^2 \times e^{1-x^2-y^2}$$

FOC:
Partial derivative in x:

$$2xe^{-x^2-y^2+1} - 2xe^{-x^2-y^2+1}\left(x^2 + y^2\right) = 0$$

Partial derivative in y:

$$2ye^{-x^2-y^2+1} - 2ye^{-x^2-y^2+1}\left(x^2 + y^2\right) = 0$$

$x^2 + y^2 = 1$ and $x = 0, y = 0$.

$$H_{11} = -8e^{-x^2-y^2+1}x^2 + 4e^{-x^2-y^2+1}\left(x^2+y^2\right)x^2 + 2e^{-x^2-y^2+1} - 2e^{-x^2-y^2+1}\left(x^2+y^2\right)$$

$$H_{12} = 4e^{-x^2-y^2+1}xy\left(x^2+y^2\right) - 8e^{-x^2-y^2+1}xy$$

$$H_{21} = 4e^{-x^2-y^2+1}xy\left(x^2+y^2\right) - 8e^{-x^2-y^2+1}xy$$

$$H_{22} = -8e^{-x^2-y^2+1}y^2 + 4e^{-x^2-y^2+1}\left(x^2+y^2\right)y^2 + 2e^{-x^2-y^2+1} - 2e^{-x^2-y^2+1}\left(x^2+y^2\right)$$

Eigenvectors of the hessian are $(-\frac{y}{x}, 1)$, $(\frac{x}{y}, 1)$.
The eigenvalues are

$$\left\{2e^{-x^2-y^2+1}\left(-x^2-y^2+1\right), 2e^{-x^2-y^2+1}\left(2x^4+4x^2y^2-5x^2+2y^4-5y^2+1\right)\right\}$$

At (0,0) the test works, but on the "rim", it is inconclusive. Eigenvalues are not all positive.
You can also perform the test through principal minors ($x^2+y^2=1$ should simplify the matrix greatly). You can check that determinant of the Hessian is zero.

$$H = \begin{bmatrix} 4\left(y^2 - 1\right) & -4xy \\ -4xy & -4y^2 \end{bmatrix}$$

Determinant is $16y^2 - 16x^2y^2 - 16y^4 = 16y^2(1 - x^2 - y^2) = 0$

With unconstrained maximization problems of the form $max\ f(x), x \in \mathbb{R}_n$, the conditions are simple (**concavity**). For constrained maximization they are lengthier.

For exact conditions that you need to check refer to the summary.

A summary of necessary and sufficient conditions for optimization problems: https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/osm/t

Part 2: Constrained optimization

Unconstrained optimization (what we did before)
find x^* that solves $max_x f(x)$.

Constrained optimization (what we are doing now):
find x^* that solves

$max_x f(x)$, s.t. $g_j(x) \leq c_j$ for $j = 1...m$

e.g.
find $x^* = (x, y)$ that solves $max\ u(x, y)$, s.t. $xp_1 + yp_2 \leq m$.

The usual approach to solution is to introduce a new variable λ, and write out a Lagrangean.

$$L(x) = f(x) - \lambda(g(x) - c)$$

$g(x) - c$ is the **slack** of the constraint. If it is zero, then the constraint **binds**.
The solution to maximization problem could be in both cases.
The intuition is that we are looking for the points where $f(x)$ and $g(x)$ are tangent, i.e. the gradients are (not equal) but are proportional and on the same line.

Let f and g_j for $j = 1, ..., m$ be differentiable functions of n variables and let c_j for $j = 1, ..., m$ be numbers. Define the function L of n variables by

$$L(x) = f(x) - \sum \lambda(g(x) - c_j)$$

for all x

The Karush-Kuhn-Tucker conditions for the problem $max_x f(x)$ subject to $g_j(x) \leq c_j$ for $j = 1, ..., m$ are

$L'i(x) = 0$ for $i = 1, ..., n$

$\lambda_j \geq 0$, $g_j(x) \leq c_j$ and $\lambda_j(g_j(x) - c_j) = 0$ for $j = 1, ..., m$.

Let's try an example from the book:

$$max_{x_1,x_2}(-(x_1-4)^2-(x_2-4)^2),$$

s.t.
$x_1+x_2 \leq 4$ and $x_1+3x_2 \leq 9$

$$L(x_1,x_2) = -(x_1-4)^2-(x_2-4)^2-\lambda_1(x_1+x_2-4)-\lambda_2(x_1+3x_2-9)$$

- KKT conditions:

Let's try an example from the book:

$$max_{x_1,x_2}(-(x_1-4)^2-(x_2-4)^2),$$

s.t.
$x_1+x_2 \leq 4$ and $x_1+3x_2 \leq 9$

$$L(x_1,x_2) = -(x_1-4)^2-(x_2-4)^2-\lambda_1(x_1+x_2-4)-\lambda_2(x_1+3x_2-9)$$

- KKT conditions:
- $-2(x_1-4)-\lambda_1-\lambda_2=0$

Let's try an example from the book:

$$max_{x_1,x_2}(-(x_1-4)^2-(x_2-4)^2),$$

s.t.
$x_1+x_2 \leq 4$ and $x_1+3x_2 \leq 9$

$$L(x_1,x_2) = -(x_1-4)^2-(x_2-4)^2-\lambda_1(x_1+x_2-4)-\lambda_2(x_1+3x_2-9)$$

- KKT conditions:
- $-2(x_1-4)-\lambda_1-\lambda_2=0$
- $-2(x_2-4)-\lambda_1-3\lambda_2=0$

Let's try an example from the book:

$$max_{x_1,x_2}(-(x_1-4)^2-(x_2-4)^2),$$

s.t.
$x_1+x_2 \leq 4$ and $x_1+3x_2 \leq 9$

$$L(x_1,x_2) = -(x_1-4)^2-(x_2-4)^2-\lambda_1(x_1+x_2-4)-\lambda_2(x_1+3x_2-9)$$

- KKT conditions:
- $-2(x_1-4)-\lambda_1-\lambda_2=0$
- $-2(x_2-4)-\lambda_1-3\lambda_2=0$
- $x_1+x_2 \leq 4$, $\lambda_1 \geq 0$ and $\lambda_1(x_1+x_2-4)=0$

Let's try an example from the book:

$$max_{x_1,x_2}(-(x_1-4)^2-(x_2-4)^2),$$

s.t.
$x_1+x_2 \leq 4$ and $x_1+3x_2 \leq 9$

$$L(x_1,x_2) = -(x_1-4)^2-(x_2-4)^2-\lambda_1(x_1+x_2-4)-\lambda_2(x_1+3x_2-9)$$

- KKT conditions:
- $-2(x_1-4)-\lambda_1-\lambda_2 = 0$
- $-2(x_2-4)-\lambda_1-3\lambda_2 = 0$
- $x_1+x_2 \leq 4$, $\lambda_1 \geq 0$ and $\lambda_1(x_1+x_2-4) = 0$
- $x_1+3x_2 \leq 9$, $\lambda_2 \geq 0$ and $\lambda_2(x_1+3x_2-9) = 0$

Is KKT Necessary or Suffient?

They can be under regularity conditions.
https://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/osm/t