Chapter 16

Kuhn Tucker Conditions

Constraint Optimization

Find the maximum of function

$$f(x,y)$$

subject to

$$g(x,y) \leq c, \qquad x,y \geq 0$$

Example:
Find the maxima of

$$f(x,y) = -(x-5)^2 - (y-5)^2$$

subject to

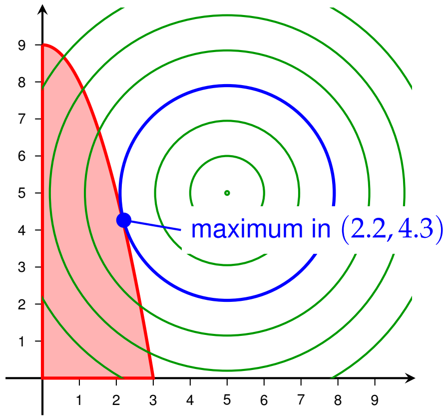
$$x^2 + y \leq 9, \qquad x,y \geq 0$$

Graphical Solution

For the case of two variables we can find a solution graphically.

1. Draw the constraint $g(x,y) \leq c$ in the xy-plain (*feasible region*).
2. Draw *appropriate* contour lines of objective function $f(x,y)$.
3. Investigate which contour lines of the objective function intersect with the feasible region.
   Estimate the (approximate) location of the maxima.

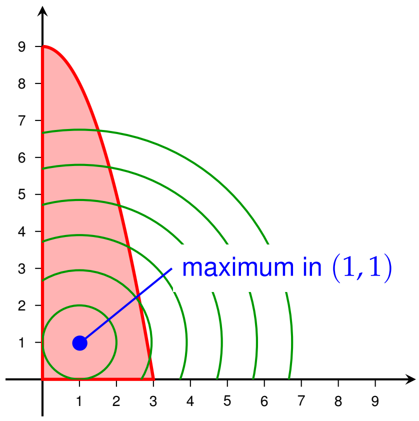
Example – Graphical Solution



Maximum of $f(x,y) = -(x-5)^2 - (y-5)^2$

subject to $g(x,y) = x^2 + y \le 9, \quad x,y \ge 0.$

Example – Graphical Solution



Maximum of $f(x,y) = -(x-1)^2 - (y-1)^2$

subject to $g(x,y) = x^2 + y \le 9, \quad x,y \ge 0.$

Constraint Optimization

Compute the maximum of function

$$f(x_1, \ldots, x_n)$$

subject to

$$\begin{aligned} g_1(x_1, \ldots, x_n) &\le c_1 \\ &\vdots \\ g_k(x_1, \ldots, x_n) &\le c_k \\ x_1, \ldots, x_n &\ge 0 \qquad \textbf{(non-negativity constraint)} \end{aligned}$$

Optimization problem:

$$\max \ f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{g}(\mathbf{x}) \le \mathbf{c} \quad \text{and} \quad \mathbf{x} \ge 0.$$

Non-Negativity Constraint

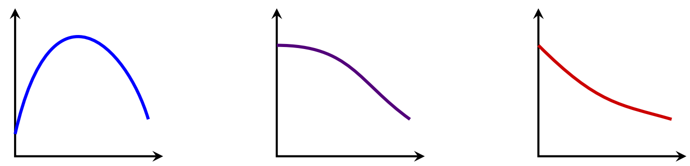
Univariate function f with non-negativity constraint.

We find for the maximum x^* of f:

- x^* is an *interior* point of the feasible region:
  $x^* > 0$ and $f'(x^*) = 0$; or
- x^* is a boundary point of the feasible region:
  $x^* = 0$ and $f'(x^*) \leq 0$.

Summary:

$$f'(x^*) \leq 0, \qquad x^* \geq 0 \qquad \text{and} \qquad x^* f'(x^*) = 0$$



Non-Negativity Constraint

For the case of a multivariate function $f(\mathbf{x})$ with non-negativity constraints $x_j \geq 0$, we obtain such a condition for each of the variables:

$$f_{x_j}(\mathbf{x}^*) \leq 0, \qquad x_j^* \geq 0 \qquad \text{and} \qquad x_j^* f_{x_j}(\mathbf{x}^*) = 0$$

Slack Variables

Maximize

$$f(x_1, \ldots, x_n)$$

subject to

$$\begin{aligned} g_1(x_1, \ldots, x_n) + \boldsymbol{s_1} &= c_1 \\ &\vdots \\ g_k(x_1, \ldots, x_n) + \boldsymbol{s_k} &= c_k \\ x_1, \ldots, x_n &\geq 0 \\ \boldsymbol{s_1}, \ldots, \boldsymbol{s_k} &\geq 0 \qquad \text{(new non-negativity constraint)} \end{aligned}$$

Lagrange function:

$$\tilde{\mathcal{L}}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) = f(x_1, \ldots, x_n) + \sum_{i=1}^{k} \lambda_i (c_i - g_i(x_1, \ldots, x_n) - \boldsymbol{s_i})$$

Slack Variables

$$\tilde{\mathcal{L}}(\mathbf{x}, \mathbf{s}, \boldsymbol{\lambda}) = f(x_1, \dots, x_n) + \sum_{i=1}^{k} \lambda_i (c_i - g_i(x_1, \dots, x_n) - s_i)$$

Apply non-negativity conditions:

$$\frac{\partial \tilde{\mathcal{L}}}{\partial x_j} \leq 0, \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial \tilde{\mathcal{L}}}{\partial x_j} = 0$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial s_i} \leq 0, \quad s_i \geq 0 \quad \text{and} \quad s_i \frac{\partial \tilde{\mathcal{L}}}{\partial s_i} = 0$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \lambda_i} = 0 \qquad \text{(no non-negativity constraint)}$$

Elimination of Slack Variables

Because of $\frac{\partial \tilde{\mathcal{L}}}{\partial s_i} = -\lambda_i$ the second line is equivalent to

$$\lambda_i \geq 0, \qquad s_i \geq 0 \qquad \text{and} \qquad \lambda_i s_i = 0$$

Equations $\frac{\partial \tilde{\mathcal{L}}}{\partial \lambda_i} = c_i - g_i(\mathbf{x}) - s_i = 0$ imply $s_i = c_i - g_i(\mathbf{x})$

and consequently the second line is equivalent to

$$\lambda_i \geq 0, \quad c_i - g_i(\mathbf{x}) \geq 0 \quad \text{and} \quad \lambda_i (c_i - g_i(\mathbf{x})) = 0\,.$$

Therefore there is no need of slack variables any more.

Elimination of Slack Variables

So we replace $\tilde{\mathcal{L}}$ by Lagrange function

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(x_1, \dots, x_n) + \sum_{i=1}^{k} \lambda_i (c_i - g_i(x_1, \dots, x_n))$$

Observe that

$$\frac{\partial \mathcal{L}}{\partial x_j} = \frac{\partial \tilde{\mathcal{L}}}{\partial x_j} \qquad \text{and} \qquad \frac{\partial \mathcal{L}}{\partial \lambda_i} = c_i - g_i(\mathbf{x})$$

So the second line of the condition for a maximum now reads

$$\lambda_i \geq 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda_i} \geq 0 \quad \text{and} \quad \lambda_i \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0$$

Kuhn-Tucker Conditions

$$\boxed{\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(x_1, \dots, x_n) + \sum_{i=1}^{k} \lambda_i (c_i - g_i(x_1, \dots, x_n))}$$

The **Kuhn-Tucker conditions** for a (global) maximum are:

$$\begin{array}{lll}
\dfrac{\partial \mathcal{L}}{\partial x_j} \leq 0, \quad x_j \geq 0 & \text{and} & x_j \dfrac{\partial \mathcal{L}}{\partial x_j} = 0 \\[2ex]
\dfrac{\partial \mathcal{L}}{\partial \lambda_i} \geq 0, \quad \lambda_i \geq 0 & \text{and} & \lambda_i \dfrac{\partial \mathcal{L}}{\partial \lambda_i} = 0
\end{array}$$

Notice that these Kuhn-Tucker conditions are not sufficient.
(Analogous to critical points.)

Example – Kuhn-Tucker Conditions

Find the maximum of

$$f(x, y) = -(x-5)^2 - (y-5)^2$$

subject to

$$x^2 + y \leq 9, \qquad x, y \geq 0$$

Example – Kuhn-Tucker Conditions

Lagrange function:

$$\mathcal{L}(x, y; \lambda) = -(x-5)^2 - (y-5)^2 + \lambda(9 - x^2 - y)$$

Kuhn-Tucker Conditions:

$$\begin{array}{rrcll}
(A) & \mathcal{L}_x & = & -2(x-5) - 2\lambda x & \leq 0 \\
(B) & \mathcal{L}_y & = & -2(y-5) - \lambda & \leq 0 \\
(C) & \mathcal{L}_\lambda & = & 9 - x^2 - y & \geq 0 \\[2ex]
(N) & & & x,\ y,\ \lambda & \geq 0 \\[2ex]
(I) & x\,\mathcal{L}_x & = & -x(2(x-5) + 2\lambda x) & = 0 \\
(II) & y\,\mathcal{L}_y & = & -y(2(y-5) + \lambda) & = 0 \\
(III) & \lambda\,\mathcal{L}_\lambda & = & \lambda(9 - x^2 - y) & = 0
\end{array}$$

Example – Kuhn-Tucker Conditions

Express equations (I)–(III) as

$$\begin{array}{lll} (I) & x=0 \quad \text{or} & 2(x-5)+2\lambda x=0 \\ (II) & y=0 \quad \text{or} & 2(y-5)+\lambda=0 \\ (III) & \lambda=0 \quad \text{or} & 9-x^2-y=0 \end{array}$$

We have to compute all 8 combinations and check whether the resulting solutions satisfy inequalities (A), (B), (C), and (N).

- If $\lambda=0$ (III,left), then by (I) and (II) there exist four solutions for $(x,y;\lambda)$:

  $(0,0;0)$, $(0,5;0)$, $(5,0;0)$, and $(5,5;0)$.

  However, none of these points satisfies all inequalities (A), (B), (C).

  Hence $\lambda \neq 0$.

Example – Kuhn-Tucker Conditions

If $\lambda \neq 0$, then (III,right) implies $y=9-x^2$.

- If $\lambda \neq 0$ and $x=0$, then $y=9$ and because of (II,right), $\lambda=-8$. A contradiction to (N).
- If $\lambda \neq 0$ and $y=0$, then $x=3$ and because of (I,right), $\lambda=\frac{2}{3}$. A contradiction to (B).
- Consequently all three variables must be non-zero.
  Thus $y=9-x^2$ and $\lambda=-2(y-5)=-2(4-x^2)$.
  Substituted in (I) yields $2(x-5)-4(4-x^2)x=0$ and

$$x=\frac{\sqrt{11}+1}{2}\approx 2.158 \qquad y=\frac{12-\sqrt{11}}{2}\approx 4.342$$
$$\lambda=\sqrt{11}-2\approx 1.317$$

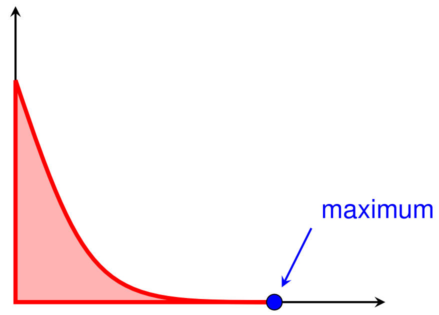
The Kuhn-Tucker conditions are thus satisfied only in point

$$(x,y;\lambda)=\left(\frac{\sqrt{11}+1}{2},\frac{12-\sqrt{11}}{2};\sqrt{11}-2\right).$$

Kuhn-Tucker Conditions

Unfortunately the Kuhn-Tucker conditions are not necessary!

That is, there exist optimization problems where the maximum does *not* satisfy the Kuhn-Tucker conditions.



Kuhn-Tucker Theorem

We need a tool to determine whether a point is a (global) maximum.

The **Kuhn-Tucker theorem** provides a *sufficient* condition:

(1) Objective function $f(\mathbf{x})$ is differentiable and **concave**.

(2) All functions $g_i(\mathbf{x})$ from the constraints are differentiable and **convex**.

(3) Point $\mathbf{x}^*$ satisfy the Kuhn-Tucker conditions.

Then $\mathbf{x}^*$ is a *global maximum* of f subject to constraints $g_i \leq c_i$.

The maximum is unique, if function f is *strictly concave*.

Example – Kuhn-Tucker Theorem

Find the maximum of

$$f(x,y) = -(x-5)^2 - (y-5)^2$$

subject to

$$x^2 + y \leq 9, \qquad x, y \geq 0$$

The respective Hessian matrices of $f(x,y)$ and $g(x,y) = x^2 + y$ are

$$\mathbf{H}_f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{and} \quad \mathbf{H}_g = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Example – Kuhn-Tucker Theorem

$$\mathbf{H}_f = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{and} \quad \mathbf{H}_g = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

(1) f is strictly concave.

(2) g is convex.

(3) Point $(x, y; \lambda) = \left(\frac{\sqrt{11}+1}{2}, \frac{12-\sqrt{11}}{2}; \sqrt{11} - 2\right)$ satisfy the Kuhn-Tucker conditions.

Thus by the Kuhn-Tucker theorem, $\mathbf{x}^* = \left(\frac{\sqrt{11}+1}{2}, \frac{12-\sqrt{11}}{2}\right)$ is the maximum we sought for.

Summary

- constraint optimization
- graphical solution
- Lagrange function
- Kuhn-Tucker conditions
- Kuhn-Tucker theorem