PRO 5961 Métodos de Otimização Não Linear

Celma de Oliveira Ribeiro

Aula 8 - Lagrangian and KKT- 2023

Departmento de Engenharia de Produção
Universidade de São Paulo

Constrained problems

The problem

Given

$x \in \mathbb{R}^n$ - variables

$f : \mathbb{R}^n \to \mathbb{R}$ - objective function

g_i e h_i constraints

P_{constr}

$$\begin{array}{lll} & \text{minimize } f(x) & \\ \text{s.t} & g_i(x) \leq 0 & \text{i} \in \{1, 2, \ldots m\} \\ & h_i(x) = 0 & \text{i} \in \{1, 2, \ldots l\} \end{array}$$

Example 0 - Budgetary constraints

Suppose you are running a factory, producing some sort of widget that requires steel as a raw material. Your costs are predominantly human labor, which is \$20 per hour for your workers, and the steel itself, which runs for \$170 per ton. Suppose your revenue is loosely modeled by the following equation: $R(h, s) = 200h^{2/3}s^{1/3}$

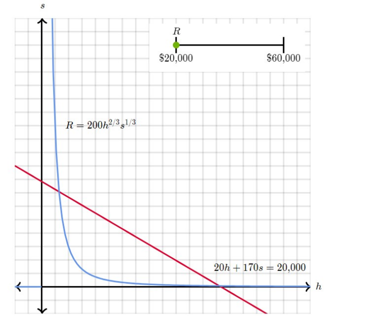
- h represents hours of labor
- s represents tons of steel

If your budget is \$20000, what is the maximum possible revenue?

The model is

$max\ 200h^{2/3}s^{1/3}$

s.t. $20h + 170s = 20,000$



Lagrange multipliers

Consider the problem P_{constr} and let $\lambda \in \mathbb{R}^l$ and $\mu \in \mathbb{R}^m$

The Lagrangian function is defined as

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^t h(x) + \mu^t g(x)$$

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x) + \sum_{i=1}^{l} \lambda_i h_i(x)$$

Append each constraint function to the objective, multiplied by a scalar for that constraint called a **Lagrange multiplier**.

Example 0 : The Lagrangian is given as:

$$\mathcal{L}(h, s, \lambda) = 200h^{2/3}s^{1/3} + \lambda(20h + 170s - 20,000)$$

What happens with critical points of $\mathcal{L}(h, s, \lambda)$? That is, when $\nabla\mathcal{L}(h, s, \lambda) = 0$?

Example 1 :

Find the points on the circle $x^2 + y^2 = 80$ which are closest to the point $(1, 2)$.

Example 1 :

Find the points on the circle $x^2 + y^2 = 80$ which are closest to the point $(1, 2)$.

$$min\ f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2$$

$$h(x_1, x_2) = x_1^2 + x_2^2 = 80$$

Example 1 :

Find the points on the circle $x^2 + y^2 = 80$ which are closest to the point $(1, 2)$.

$$min\ f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 2)^2$$

$$h(x_1, x_2) = x_1^2 + x_2^2 = 80$$

Lagrangian:

$$\mathcal{L}(x_1, x_2, \lambda) = (x_1 - 1)^2 + (x_2 - 2)^2 + \lambda(x_1^2 + x_2^2 - 80)$$

l = 1 and n = 2

Example 2 :

$$minf(x) = x_1^2 - x_2^2$$

$$x_1 + 2x_2 + 1 = 0$$
$$x_1 - x_2 \leq 3$$

Example 2 :

$$minf(x) = x_1^2 - x_2^2$$

$$x_1 + 2x_2 + 1 = 0$$
$$x_1 - x_2 \leq 3$$

Lagrangian:

$$\mathcal{L}(x_1, x_2, \lambda, \mu) = x_1^2 - x_2^2 + \lambda(x_1 + 2x_2 + 1) + \mu\,(x_1 - x_2 - 3)$$

$m = 1$, $l = 1$ and $n = 2$

What happens if we consider the problem $\min_{x\lambda,\mu} \mathcal{L}(x\lambda, \mu)$?

What happens if we consider the problem $\min_{x\lambda,\mu} \mathcal{L}(x\lambda,\mu)$?

Example 3 : Consider

$$\min f(x) = -4x_1 + 0.1x_1^2 - 5x_2 + 0.2x_2^2$$

$$x_1 + 2x_2 = 40$$

Build the Lagrangian and its gradient

Lagrange multipliers

What happens if we consider the problem $\min_{x\lambda,\mu} \mathcal{L}(x\lambda,\mu)$?

Example 3 : Consider

$$\min f(x) = -4x_1 + 0.1x_1^2 - 5x_2 + 0.2x_2^2$$
$$x_1 + 2x_2 = 40$$

Build the Lagrangian and its gradient

Lagrangian $\mathcal{L}(x,\lambda) = -4x_1 + 0.1x_1^2 - 5x_2 + 0.2x_2^2 + \lambda_1(x_1 + 2x_2 - 40)$

What happens if we consider the problem $\min_{x\lambda,\mu} \mathcal{L}(x\lambda, \mu)$?

Example 3 : Consider

$$\min f(x) = -4x_1 + 0.1x_1^2 - 5x_2 + 0.2x_2^2$$
$$x_1 + 2x_2 = 40$$

Build the Lagrangian and its gradient

Lagrangian $\mathcal{L}(x, \lambda) = -4x_1 + 0.1x_1^2 - 5x_2 + 0.2x_2^2 + \lambda_1(x_1 + 2x_2 - 40)$

Gradient

$$\nabla\mathcal{L}(x, \lambda) = \begin{bmatrix} -4 + 0.2x_1 + \lambda_1 \\ -5 + 0.4x_2 + 2\lambda_1 \\ x_1 + 2x_2 - 40 \end{bmatrix}$$

Method of Lagrange multipliers

What happens when we analyze the gradient of the Lagrangian?

$$\nabla \mathcal{L}\left(x^{*}, \lambda^{*}\right)=0$$

Method of Lagrange multipliers

What happens when we analyze the gradient of the Lagrangian?

$$\nabla \mathcal{L}\left(x^*, \lambda^*\right) = 0$$

$$\nabla \mathcal{L}\left(x, \lambda\right) = \begin{bmatrix} -4 - +0.2x_1 + \lambda_1 \\ -5 + 0.4x_2 + 2\lambda_1 \\ x_1 + 2x_2 - 40 \end{bmatrix} = 0$$

Solving the system

$$\nabla \mathcal{L}\left(x^*, \lambda^*\right) = 0$$

we obtain

$x_1^* \approx 18.3$, $x_2^* \approx 10.8$ and $\lambda^* = 0.33$

Which is the optimal solution of the original problem

Once we have found candidate solutions x^*, it is not always easy to figure out whether it corresponds to a minimum, a maximum or neither.

Interpretation of Lagrange multipliers

Consider in Example 0, λ^* such that

$$\nabla L(h, s, \lambda^*) = 0$$

λ^* tells us how much more money we can make by changing our budget.

Observe that the partial derivatives of the Lagrangian

$$\mathcal{L}(x,\lambda) = f(x) + \lambda^t h(x) + \mu^t g(x)$$

are as follows:

$$\nabla_x \mathcal{L}(x,\lambda) = \nabla_x f(x) + \lambda^t \nabla_x h(x) + \mu^t \nabla_x g(x)$$

$$\nabla_\lambda \mathcal{L}(x,\lambda) = h(x)$$

$$\nabla_\mu \mathcal{L}(x,\lambda) = g(x)$$

Exercice: analyze what happens in Example 1

Entrega aula For example 2, verify the conditions above

Specific situations shall be addressed. Equality constraints are different from inequalities...

Consider P_{equal}

$$\begin{array}{lll} & \text{minimize } f(x) & \\ \text{s.t} & h_i(x) = 0 & \text{i} \in \{1, 2, \ldots l\} \end{array}$$

Let us look for the points $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n, \bar{\lambda}_1, \bar{\lambda}_2, ...\bar{\lambda}_l)$ for which

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial \mathcal{L}}{\partial x_2} = ... = \frac{\partial \mathcal{L}}{\partial x_n} = \frac{\partial \mathcal{L}}{\partial \lambda_1} = \frac{\partial \mathcal{L}}{\partial \lambda_2} = ... = \frac{\partial \mathcal{L}}{\partial \lambda_l} = 0 \qquad (1)$$

In many situations $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n, \bar{\lambda}_1, \bar{\lambda}_2, ...\bar{\lambda}_l)$ solves the original problem

Theorem - Important

If $f(x)$ is a convex function and $h_i(x)$ is linear for all i, then any point $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n, \bar{\lambda}_1, \bar{\lambda}_2, ...\bar{\lambda}_l)$ satisfying (1) will yield an optimal solution $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_n)$ to P_{equal}

Analyze the previous examples

Example

$\min 6x_1^2 + 4x_2^2 + 3x_3^2$

$24x_1 + 24x_2 = 360$

$x_3 = 1$

Lagrangian

Example

$\min 6x_1^2 + 4x_2^2 + 3x_3^2$

$24x_1 + 24x_2 = 360$

$x_3 = 1$

$\mathcal{L}(x, \lambda) = 6x_1^2 + 4x_2^2 + 3x_3^2 + \lambda_1 (360 - 24x_1 - 24x_2) + \lambda_2 (1 - x_3)$

$$\nabla f(x)^t = \begin{bmatrix} 12x_1 & 8x_2 & 6x_3 \end{bmatrix}$$

$$\nabla h_1(x)^t = \begin{bmatrix} -24 & -24 & 0 \end{bmatrix}$$

$$\nabla h_2(x)^t = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

- $\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) + \lambda_1 \nabla h_1(x) + \lambda_2 \nabla h_2(x) = 0$

  $\boxed{\text{Linear system}}$

Lagrangian

- $\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) + \lambda_1 \nabla h_1(x) + \lambda_2 \nabla h_2(x) = 0$

  Linear system

  $$\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow \begin{cases} 12x_1 & = 24\lambda_1 \\ 8x_2 & = 24\lambda_1 \\ 6x_3 & = 1\lambda_2 \end{cases}$$

- $\nabla_\lambda \mathcal{L}(x, \lambda) = 0 \Rightarrow h(x) = 0 \Rightarrow \begin{cases} 24x_1 + 24x_2 & = 360 \\ x_3 & = 1 \end{cases}$

Solution of the system: $(x^*, \lambda^*)^t = (2, 3, 1, 9, 6)$

$f(x)$ is a convex function and $h_i(x)$ is linear for all i, $\Rightarrow x^*$ is a solution of the problem

Lagrangian

- $\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) + \lambda_1 \nabla h_1(x) + \lambda_2 \nabla h_2(x) = 0$

  $\boxed{\text{Linear system}}$

  $$\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow \begin{cases} 12x_1 & = 24\lambda_1 \\ 8x_2 & = 24\lambda_1 \\ 6x_3 & = 1\lambda_2 \end{cases}$$

- $\nabla_\lambda \mathcal{L}(x, \lambda) = 0 \Rightarrow h(x) = 0 \Rightarrow \begin{cases} 24x_1 + 24x_2 & = 360 \\ x_3 & = 1 \end{cases}$

Solution of the system: $(x^*, \lambda^*)^t = (2, 3, 1, 9, 6)$

$f(x)$ is a convex function and $h_i(x)$ is linear for all i, $\Rightarrow x^*$ is a solution of the problem

$\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow -\nabla f(x) = \lambda_1 \nabla h_1(x) + \lambda_2 \nabla h_2(x)$

Entrega semanal

A company is planning to spend \$10 on advertising. It costs \$3 per minute to advertise on TV and \$1 per minute to advertise on radio.If the firm buys x minutes of TV advertising and y minutes of radio advertising, its revenue is given as

$$f(x, y) = -2x^2 - y^2 + xy + 8x + 3y$$

. How can the firm maximize its revenue?

Hint:

- Write the optimization problem (in the min form) and the lagragian
- Find the partial derivatives and set them $= 0$
- Solve the resulting system
- Analyse the Hessian

Lagrangian

Some important questions:

- Can the solution of (P_{equal}) be obtained through unconstrained optimization considering the Lagrangian as a *penalty* function?

Some important questions:

- Can the solution of (P_{equal}) be obtained through unconstrained optimization considering the Lagrangian as a *penalty* function?
- When critical points(w^* is a critical point of $\psi(w)$ if $\nabla\psi(w^*) = 0$) of the Lagrangian are optimal solutions of the original problem?

Some important questions:

- Can the solution of (P_{equal}) be obtained through unconstrained optimization considering the Lagrangian as a *penalty* function?
- When critical points(w^* is a critical point of $\psi(w)$ if $\nabla\psi(w^*) = 0$) of the Lagrangian are optimal solutions of the original problem?
- Does an optimal solution of the problem provide a critical point for the Lagrangian?

Lagrangian

Does an optimal solution of the problem provide a critical point for the Lagrangian?

Consider

min $x_1 + x_2 + x_3^2$

$$\text{s.t.} \quad \begin{array}{l} x_1 = 1 \\ x_1^2 + x_2^2 = 1 \end{array}$$

- The minimum is achieved at $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^t$
- The associated Lagrangian is:

Lagrangian

Does an optimal solution of the problem provide a critical point for the Lagrangian?

Consider

min $x_1 + x_2 + x_3^2$

s.t. $\begin{array}{l} x_1 = 1 \\ x_1^2 + x_2^2 = 1 \end{array}$

- The minimum is achieved at $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^t$
- The associated Lagrangian is: $\mathcal{L}(x, \lambda) = x_1 + x_2 + x_3^2 + \lambda_1(1 - x_1) + \lambda_2(1 - x_1^2 - x_2^2 :)$

Lagrangian

Does an optimal solution of the problem provide a critical point for the Lagrangian?

Consider

$$\min\ x_1 + x_2 + x_3^2$$

$$\text{s.t.}\quad \begin{array}{l} x_1 = 1 \\ x_1^2 + x_2^2 = 1 \end{array}$$

- The minimum is achieved at $\left[\begin{array}{ccc} 1 & 0 & 0 \end{array}\right]^t$
- The associated Lagrangian is:
  $\mathcal{L}\left(x,\lambda\right) = x_1 + x_2 + x_3^2 + \lambda_1(1 - x_1) + \lambda_2(1 - x_1^2 - x_2^2 :)$
- Write $\dfrac{\partial \mathcal{L}}{\partial x_2}$ and explain what happens at $\left[\begin{array}{ccccc} 1 & 0 & 0 & \lambda_1 & \lambda_2 \end{array}\right]^t$

Lagrangian

Does an optimal solution of the problem provide a critical point for the Lagrangian?

Consider

min $x_1 + x_2 + x_3^2$

s.t. $\begin{array}{l} x_1 = 1 \\ x_1^2 + x_2^2 = 1 \end{array}$

- The minimum is achieved at $\left[\begin{array}{ccc} 1 & 0 & 0 \end{array}\right]^t$
- The associated Lagrangian is:
  $\mathcal{L}(x, \lambda) = x_1 + x_2 + x_3^2 + \lambda_1(1 - x_1) + \lambda_2(1 - x_1^2 - x_2^2 :)$
- Write $\dfrac{\partial \mathcal{L}}{\partial x_2}$ and explain what happens at $\left[\begin{array}{ccccc} 1 & 0 & 0 & \lambda_1 & \lambda_2 \end{array}\right]^t$
  $\dfrac{\partial \mathcal{L}}{\partial x_2}(1, 0, 0, \lambda_1, \lambda_2) = 1,\ \forall \lambda_1, \lambda_2$ and it does not vanish at the optimal solution
  It means that $\nabla \mathcal{L}(1, 0, 0, \lambda_1, \lambda_2) \neq 0$

Entrega semanal: Write $\nabla h_1(x)$ and $\nabla h_2(x)$. Verify that these vectors are linearly dependent for $\bar{x}^t = \left[\begin{array}{ccc} 1 & 0 & 0 \end{array}\right]$

Example 4

Consider

$$\min\ 2x_1^2 + x_2^2$$

$$\text{s.t.} \quad x_1 + x_2 = 1$$

Define the Lagrangian

Example 4

Consider

min $2x_1^2 + x_2^2$

s.t. $x_1 + x_2 = 1$

Define the Lagrangian $\mathcal{L}(x, \lambda) = 2x_1^2 + x_2^2 + \lambda_1(1 - x_1 - x_2)$

Example 4

Consider

min $2x_1^2 + x_2^2$

s.t. $\quad x_1 + x_2 = 1$

Define the Lagrangian $\mathcal{L}(x, \lambda) = 2x_1^2 + x_2^2 + \lambda_1(1 - x_1 - x_2)$

Build the gradient of the Lagrangian

Example 4

Consider

$\min\ 2x_1^2 + x_2^2$

s.t. $\quad x_1 + x_2 = 1$

Define the Lagrangian $\mathcal{L}(x, \lambda) = 2x_1^2 + x_2^2 + \lambda_1(1 - x_1 - x_2)$

Build the gradient of the Lagrangian

$$\nabla \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} 4x_1^* - \lambda_1^* \\ 2x_2^* - \lambda_1^* \\ 1 - x_1^* - x_2^* \end{bmatrix}$$

Example 4

Consider

min $2x_1^2 + x_2^2$

s.t. $\quad x_1 + x_2 = 1$

Define the Lagrangian $\mathcal{L}(x, \lambda) = 2x_1^2 + x_2^2 + \lambda_1(1 - x_1 - x_2)$

Build the gradient of the Lagrangian

$$\nabla\mathcal{L}(x^*, \lambda^*) = \left[\begin{array}{l} 4x_1^* - \lambda_1^* \\ 2x_2^* - \lambda_1^* \\ 1 - x_1^* - x_2^* \end{array} \right] = 0 \Rightarrow$$

$$x_1^* = \frac{1}{3}, x_2^* = \frac{2}{3}, \lambda_1^* = \frac{4}{3}$$

The above solution is optimal

Theorem

Consider

Assume x^* is an optimal solution of

$$\min_{x \in \mathbb{R}^n} \{f(x) | h(x) = 0\}$$

Then either

i the vectors $\nabla h_1(x^*), \nabla h_2(x^*), \ldots \nabla h_l(x^*)$ are linearly dependent, or

ii there exists a vector λ^* such that $\nabla \mathcal{L}(x^*, \lambda^*) = 0$

Notation: $\nabla_x \mathcal{L}$ and $\nabla_\lambda \mathcal{L}$

Usually we cannot assure that optimal solutions are critical points of the Lagrangian!

These are **necessary** conditions for optimality

Recapitulation

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ and consider the problem

$$\min_{x \in \mathbb{R}^n} \{f(x) | g(x) \leq 0, h(x) = 0\}$$

Definition

Consider $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}$ defined by

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \mu^t g(x) + \lambda^t h(x) \forall x \in \mathbb{R}^n, \mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^l$$

The function $\mathcal{L}$ is the **Lagrangian** and the variables λ and μ are the **dual variables**

Main ideia: Find (x, λ, μ) such that $\nabla \mathcal{L}(x, \lambda, \mu) = 0$

Optimality conditions for constrained problems

Main idea

$\boxed{\text{Algebraic characterizations of solutions allowing computations.}}$

- **Sufficient conditions** provide a way to guarantee that a candidate point is optimal
- **Necessary conditions** indicate when a point is *not* optimal

Optimality conditions for constrained problems

Main idea

Algebraic characterizations of solutions allowing computations.

Sufficient conditions provide a way to guarantee that a candidate point is optimal

Necessary conditions indicate when a point is *not* optimal

Unconstrained problems $\Rightarrow$ Analyze x^*, a stationary point $(\nabla f(x^*) = 0)$

Constrained problems $\Rightarrow$ Analyze x^*,a Karush-Kuhn-Tucker (KKT) point

Example - equality

Consider

$$\min f(x) = x_1 + x_2$$

$$h(x) = x_1^2 + x_2^2 - 2 = 0$$

The unique solution is given by $x^* = [-1 - 1]^T$

Computing the gradients of f and h in x^*,

Example - equality

Consider

$$\min f(x) = x_1 + x_2$$
$$h(x) = x_1^2 + x_2^2 - 2 = 0$$

The unique solution is given by $x^* = [-1 - 1]^T$

Computing the gradients of f and h in x^*,

$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\nabla h(x^*) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

Example - equality

Consider

$$\min f(x) = x_1 + x_2$$

$$h(x) = x_1^2 + x_2^2 - 2 = 0$$

The unique solution is given by $x^* = [-1 - 1]^T$

Computing the gradients of f and h in x^*,

$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\nabla h(x^*) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

$\nabla f(x^*)$ and $\nabla h(x^*)$ are parallel, i.e., there exists a scalar $\lambda = \frac{1}{2}$ such that

$$-\nabla f(x^*) = \lambda \nabla h(x^*)$$

Example - equality

Consider

$$\min f(x) = x_1 + x_2$$

$$h(x) = x_1^2 + x_2^2 - 2 = 0$$

The unique solution is given by $x^* = [-1 - 1]^T$

Computing the gradients of f and h in x^*,

$\nabla f(x^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\nabla h(x^*) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

$\nabla f(x^*)$ and $\nabla h(x^*)$ are parallel, i.e., there exists a scalar $\lambda = \frac{1}{2}$ such that

$$-\nabla f(x^*) = \lambda \nabla h(x^*)$$

This is a necessary condition for optimality in the general case

Consider the following example

$$\min f(x) = 2x_1^2 + x_2^2$$

$$h(x) = x_1 + x_2 = 1$$

The Lagrangian is: $L(x_1, x_2, \lambda) = 2x_1^2 + x_2^2 + \lambda(1 - x_1 - x_2)$

Solve for the following:

$$\begin{cases} \nabla_x L = 0 \\ \nabla_\lambda L = 0 \end{cases}$$

Consider the following example

$$\min f(x) = 2x_1^2 + x_2^2$$

$$h(x) = x_1 + x_2 = 1$$

The Lagrangian is: $L(x_1, x_2, \lambda) = 2x_1^2 + x_2^2 + \lambda(1 - x_1 - x_2)$

Solve for the following:

$$\begin{cases} \nabla_x L = 0 \\ \nabla_\lambda L = 0 \end{cases}$$

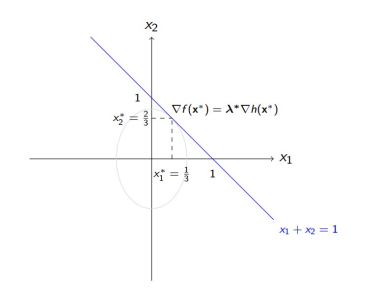
Solving this system of equations yields

$$x_1^* = \frac{1}{3}, x_2^* = \frac{2}{3}, \lambda^* = \frac{4}{3}$$

Is this a minimum or a maximum?

Optimality conditions

Graphically



- Consider the gradients of f and h at the optimal point
- They must point in the same direction, though they may have different lengths $\nabla f(x^*) = \lambda \nabla h(x^*)$

Consider problem

$$\begin{array}{lll} & \text{minimize } f(x) & \\ \text{s.t} & g_i(x) \leq 0 & \text{i} \in \mathcal{I} = \{1, 2, \ldots m\} \\ & h_i(x) = 0 & \text{i} \in \mathcal{E} = \{1, 2, \ldots l\} \end{array}$$

First order necessary Conditions

x^* is KKT point if there are lagrange multipliers vectors λ^* and μ^*, such that $\left[\begin{array}{ccc} x^* & \lambda^* & \mu^* \end{array}\right]^t$ satisfies:

$$\begin{array}{lll} \nabla_x \mathcal{L}\left(x^*, \lambda^* \mu^*\right) & = & 0 \\ g\left(x^*\right) & \leq 0 & \\ h\left(x^*\right) & = 0 & \\ \mu^* & \geq 0 & \\ \mu_i g_i(x^*) & = 0 & \forall i \in \mathcal{I} \end{array}$$

Karush-Kuhn–Tucker Conditions

Example

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

$x + 4y \leq 3$

$y \leq x$

Optimality conditions

Example

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

$x + 4y \leq 3$

$y \leq x$

Lagrangian

$\mathcal{L}(x, y, \mu_1, \mu_2) = (x-2)^2 + 2(y-1)^2 + \mu_1(x + 4y - 3) + \mu_2(-x + y)$

KKT conditions

Example

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

$x + 4y \leq 3$

$y \leq x$

Lagrangian

$\mathcal{L}(x, y, \mu_1, \mu_2) = (x-2)^2 + 2(y-1)^2 + \mu_1(x + 4y - 3) + \mu_2(-x + y)$

KKT conditions

- $\nabla_x \mathcal{L}(x, y, \mu_1, \mu_2) == 2(x-2) + \mu_1 - \mu_2 = 0$
- $\nabla_y \mathcal{L}(x, y, \mu_1, \mu_2) = 4(y-1) + 4\mu_1 + \mu_2 = 0$
- $x + 4y - 3 \leq 0$
- $y - x \leq 0$
- $\mu_1(x + 4y - 3) = 0$
- $\mu_2(y - x) = 0$
- $\mu_1, \mu_2 \geq 0$

Optimality conditions

Check 4 cases

1. $\mu_1 = \mu_2 = 0 \Rightarrow x = 2, y = 1$
2. $\mu_1 = 0 \ y - x = 0 \Rightarrow x = \frac{4}{3}, \mu_2 = -\frac{4}{3}$
3. $\mu_2 = 0 \ x + 4y - 3 = 0 \Rightarrow x = \frac{5}{3}, y = \frac{1}{3}, \mu_1 = \frac{2}{3}$
4. $x + 4y - 3 = 0 \ y - x = 0 \Rightarrow x = \frac{3}{5}, y = \frac{3}{5}, \mu_1 = \frac{22}{25}, \mu_2 = -\frac{48}{25}$

Optimal solution: $x = \frac{4}{3}$, $y = \frac{4}{3}$, $f(x, y) = \frac{4}{9}$

Does it always happen?