Constrained Optimization

Paweł Polak

November 13, 2017

Linear Regression Models - Lecture 13

Optimization

- We want to find the maximum or minimum of a function subject to some constraints.
- Given functions

$$f, g_1, \ldots, g_m \text{ and } h_1, \ldots, h_l$$

defined on some domain $\Omega \subset \mathbb{R}^n$ the optimization problem has the form

$$\min_{x \in \Omega} f(x)$$

subject to

$$g_i(x) \leq 0 \text{ for all } i = 1, \ldots, m \text{ and } h_j(x) = 0 \text{ for all } j = 1, \ldots, l$$

We will derive/state sufficient and necessary conditions for (local) optimality when there are

1. no constraints,
2. only equality constraints
3. only inequality constraints
4. equality and inequality constraints - homework

Unconstrained Optimization

Let $f : \Omega \to \mathbb{R}$ be a continuously differentiable function.

Necessary and sufficient conditions for a local minimum: x^* is local optimum of $f(x)$ if and only if

- f has zero gradient at x^*

$$\nabla_x f(x^*) = 0$$

- and the Hessian of f at x^* is

(min) positive semi-definite

$$v^T \nabla_x^2 f(x^*) v \geq 0 \text{ for all } v \in \mathbb{R}^n$$

(max) negative semi-definite

$$v^T \nabla_x^2 f(x^*) v \leq 0 \text{ for all } v \in \mathbb{R}^n$$

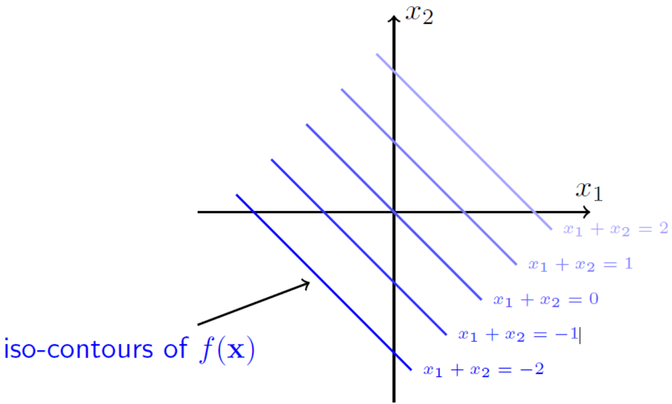
where $\nabla_x^2 f(x^*) = [\frac{\partial^2 f(x)}{\partial x_i \partial x_j}]_{i,j=1,\ldots,n}$

Constrained Optimization: Equality Constraints

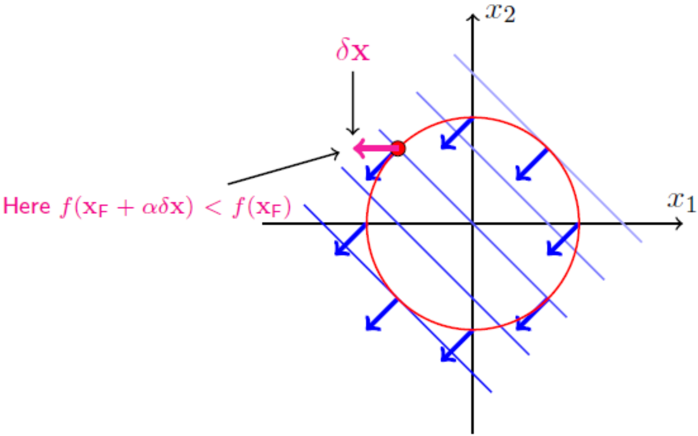
$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } h(x) = 0$$

where

$$f(x) = x_1 + x_2 \text{ and } h(x) = x_1^2 + x_2^2 - 2$$



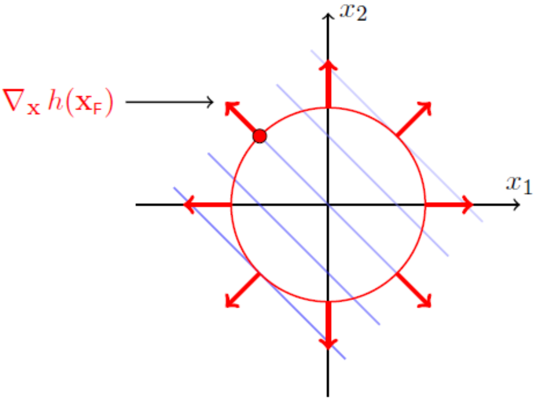
Constrained Optimization: Equality Constraints



To move δx from x such that $f(x + \delta x) < f(x)$ must have

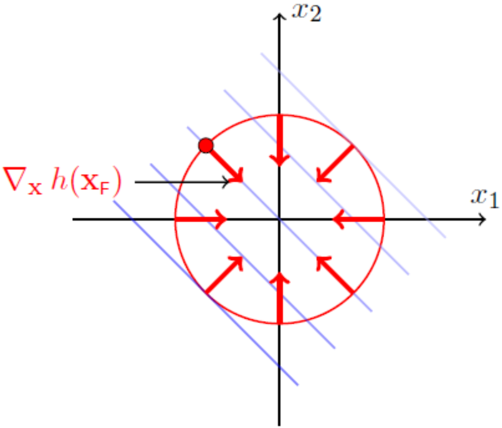
$$\delta x(-\nabla_x f(x)) > 0$$

Constrained Optimization: Equality Constraints



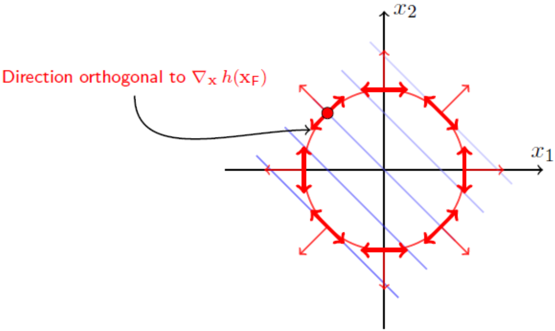
Normals to the constraint surface are given by $\nabla_x h(x)$

Constrained Optimization: Equality Constraints



Note the direction of the normal is arbitrary as the constraint be imposed as either $h(x) = 0$ or $-h(x) = 0$.

Constrained Optimization: Equality Constraints



To move a small δx from x and remain on the constraint surface we have to move in a direction orthogonal to $\nabla_x h(x)$.
If x_F lies on the constraint surface:

- setting δx orthogonal to $\nabla_x h(x_F)$ ensures $h(x_F + \delta x) = 0$ and
- $f(x_F + \delta x) < f(x_F)$ only if

$$\delta x(-\nabla_x f(x_F)) > 0.$$

Constrained Optimization: Equality Constraints

Consider the case

$$\nabla_x f(x_F) = \mu \nabla_x h(x_F),$$

where μ is a scalar.

When this occurs

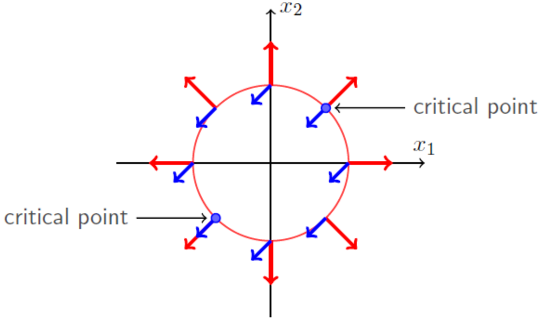
- if δx is orthogonal to $\nabla_x h(x_F)$ then

$$\delta x(-\nabla_x f(x_F)) = -\delta x \mu \nabla_x h(x_F) = 0$$

- cannot move from x_F to remain on the constraint surface and decrease (or increase) the cost function.

This case corresponds to constrained local optimum.

Constrained Optimization: Equality Constraints



A constraint local optimum occurs at x^* when $\nabla_x f(x^*)$ and $\nabla_x h(x^*)$ are parallel, i.e.,

$$\nabla_x f(x^*) = \mu \nabla_x h(x^*).$$

Constrained Optimization: Equality Constraints

We can replace our constrained optimization problem

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } h(x) = 0$$

by the Lagrangian, which is defined by

$$\mathcal{L}(x, \mu) = f(x) + \mu h(x)$$

Then the local minimum $\Leftrightarrow$ there exists a unique μ^* s.t.

- $\nabla_x \mathcal{L}(x^*, \mu^*) = 0$
- $\nabla_\mu \mathcal{L}(x^*, \mu^*) = 0$
- $y^T \nabla^2_{xx} \mathcal{L}(x^*, \mu^*) y \geq 0$ for all y s.t. $\nabla_x h(x^*)^T y = 0$.

Constrained Optimization: Equality Constraints

We can replace our constrained optimization problem

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } h(x) = 0$$

by the Lagrangian, which is defined by

$$\mathcal{L}(x, \mu) = f(x) + \mu h(x) \text{ note } \mathcal{L}(x^*, \mu^*) = f(x^*)$$

Then the local minimum $\Leftrightarrow$ there exists a unique μ^* s.t.

- $\nabla_x \mathcal{L}(x^*, \mu^*) = 0$ encodes $\nabla_x f(x^*) = \mu^* \nabla_x h(x^*)$
- $\nabla_\mu \mathcal{L}(x^*, \mu^*) = 0$ encodes the equality constraint $h(x^*) = 0$
- $y^T \nabla^2_{xx} \mathcal{L}(x^*, \mu^*) y \geq 0$ for all y s.t. $\nabla_x h(x^*)^T y = 0$ (semi-positive definite Hessian tells us that we have a local minimum).

Constrained Optimization: Equality Constraints

The general constrained optimization problem is

$$\min_{x \in \mathbb{R}^2} f(x) \text{ subject to } h_i(x) = \text{ for } i = 1, \ldots, l$$

Construct the Lagrangian (introduce a multiplier for each constraint)

$$\mathcal{L}(x, \mu) = f(x) + \sum_{i=1}^{l} \mu_i h_i(x) = f(x) + \mu^T h(x)$$

Then x^* is a local minimum if and only if there exists a uniques μ^* such that

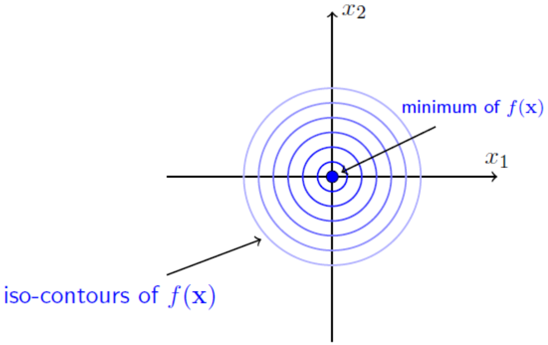
1. $\nabla_x \mathcal{L}(x^*, \mu^*) = 0$
2. $\nabla_\mu \mathcal{L}(x^*, \mu^*) = 0$
3. $y^T \nabla^2_{xx} \mathcal{L}(x^*, \mu^*) y \geq 0$ for all y such that $\nabla_x h(x^*)^T y = 0$

Constrained Optimization: Inequality Constraints

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

where

$$f(x) = x_1^2 + x_2^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$

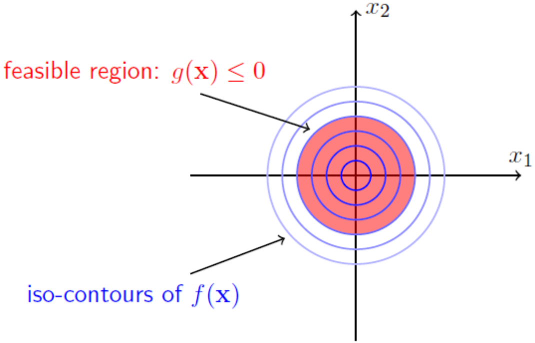


Constrained Optimization: Inequality Constraints

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

where

$$f(x) = x_1^2 + x_2^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$



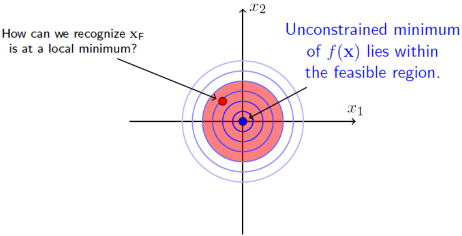
Constrained Optimization: Inequality Constraints

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

where

$$f(x) = x_1^2 + x_2^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$

How do we recognize if x_F is at a local optimum?



Easy in this case: Necessary and sufficient conditions for a constrained local minimum are the same as for an unconstrained local minimum.

$$\nabla_x f(x_F) = 0 \text{ and } \nabla_{xx}^2 f(x_F) \text{ is positive definite}$$

Constrained Optimization: Inequality Constraints

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

where

$$f(x) = x_1^2 + x_2^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$

Constraint is not active at the local minimum $g(x^*) < 0$. Therefore, the local minimum is identified by the same conditions as in the unconstrained case.

What if the constraint is inactive?

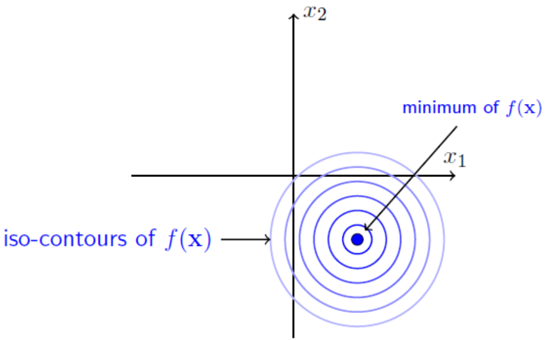
Constrained Optimization: Inequality Constraints

Suppose now that this is a constrained optimization problem which we want to solve:

$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

where

$$f(x) = (x_1 - 1.1)^2 + (x_2 + 1.1)^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$

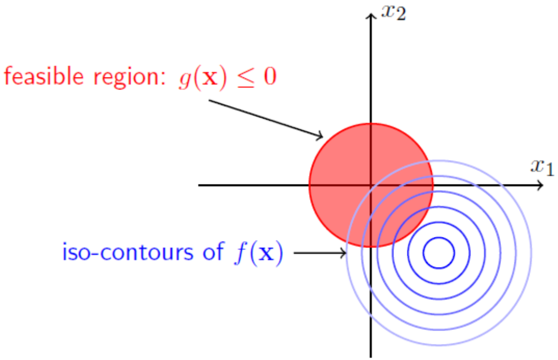


Constrained Optimization: Inequality Constraints

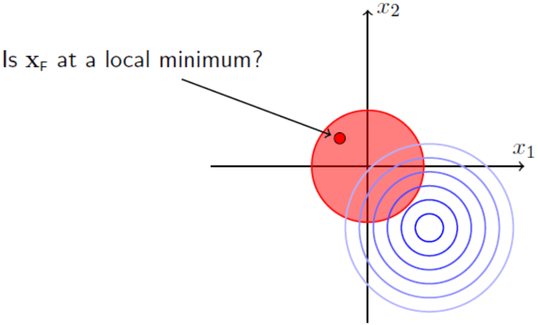
$$min_{x \in \mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

where

$$f(x) = (x_1 - 1.1)^2 + (x_2 + 1.1)^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$



Constrained Optimization: Inequality Constraints



Remember $\mathbf{x}_F$ denotes a feasible point.

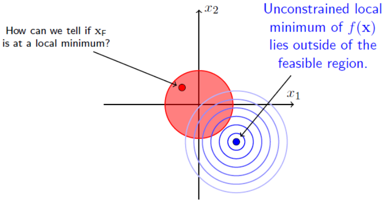
How do we recognize if x_F is at a local optimum?
Remember x_F denotes a feasible point.

Constrained Optimization: Inequality Constraints

$$min_{x\in\mathbb{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

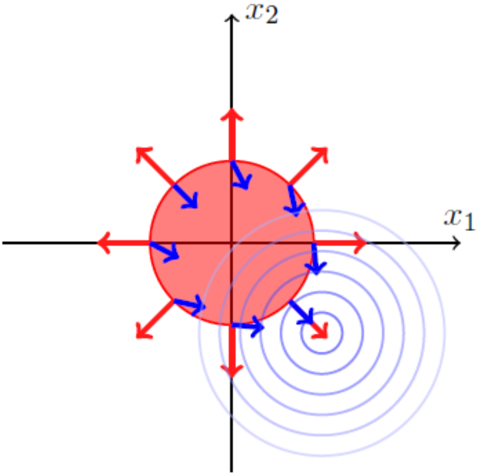
where

$$f(x) = (x_1 - 1.1)^2 + (x_2 + 1.1)^2 \text{ and } h(x) = x_1^2 + x_2^2 - 1$$



- The constrained local minimum occurs on the surface of the constraint surface.
- Effectively we have an optimization problem with an equality constraint: $g(x) = 0$.

Constrained Optimization: Inequality Constraints

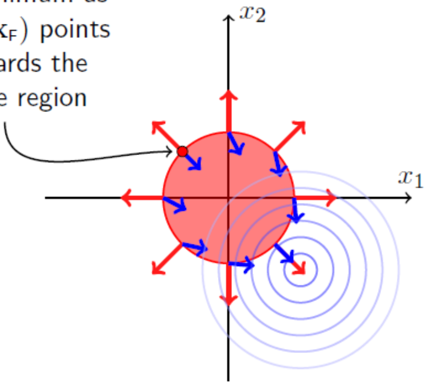


A local optimum occurs when $\nabla_x f(x)$ and $\nabla_x g(x)$ are parallel:

$$-\nabla_x f(x) = \lambda \nabla_x g(x)$$

Constrained Optimization: Inequality Constraints

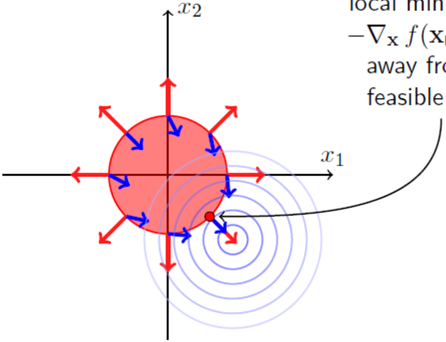
✗ **Not** a constrained local minimum as $-\nabla_{\mathbf{x}} f(\mathbf{x}_F)$ points in towards the feasible region



Constrained local minimum occurs when $-\nabla_x f(x)$ and $\nabla_x g(x)$ point in the same direction:

$$-\nabla_x f(x) = \lambda \nabla_x g(x) \text{ and } \lambda > 0$$

Constrained Optimization: Inequality Constraints



✓ **Is** a constrained local minimum as $-\nabla_{\mathbf{x}} f(\mathbf{x}_F)$ points away from the feasible region

Constrained local minimum occurs when $-\nabla_x f(x)$ and $\nabla_x g(x)$ point in the same direction:

$$-\nabla_x f(x) = \lambda \nabla_x g(x) \text{ and } \lambda > 0$$

Summary of optimization with one inequality constraint

Given

$$\min_{x \in \mathcal{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

If x^* corresponds to a constrained local minimum then,

- Case 1: Unconstrained local minimum occurs in the feasible region.
  - $g(x^*) < 0$
  - $\nabla_x f(x^*) = 0$
  - $\nabla_{xx} f(x^*)$ is positive semi-definite matrix.
- Case 2: Unconstrained local minimum lies outside the feasible region.
  - $g(x^*) = 0$
  - $-\nabla_x f(x^*) = \lambda \nabla_x g(x^*)$ with $\lambda > 0$
  - $\nabla_{xx} f(x^*)$ is positive semi-definite matrix for all y orthogonal to $\nabla_x g(x^*)$.

Karush - Kuhn -Tucker conditions encode these conditions

Given

$$\min_{x \in \mathcal{R}^2} f(x) \text{ subject to } g(x) \leq 0$$

Define the Lagrangian

$$\mathcal{L}(x, \lambda) = f(x) + \lambda g(x)$$

Then x^* corresponds to a constrained local minimum if and only if there exists a unique λ^* such that

- $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$
- $\lambda^* \geq 0$
- $\lambda^* g(x^*) = 0$
- $g(x^*) \leq 0$
- plus positive definite constraints on $\nabla_{xx} \mathcal{L}(x^*, \mu^*)$

These are the KKT conditions.

What the KKT conditions imply

- Case 1: Inactive constraint:
  - When $\lambda^* = 0$ then we have $\mathcal{L}(x^*, \mu^*) = f(x^*)$
  - KKT 1 $\Rightarrow \nabla_x f(x^*) = 0$
  - KKT 4 $\Rightarrow x^*$ is a feasible point
- Case 2: Active constraint:
  - When $\lambda^* > 0$ then we have $\mathcal{L}(x^*, \mu^*) = f(x^*) + \lambda^* g(x^*)$
  - KKT 1 $\Rightarrow \nabla_x f(x^*) = -\lambda^* g(x^*)$
  - KKT 3 $\Rightarrow g(x^*) = 0$
  - KKT 3 also $\Rightarrow \mathcal{L}(x^*, \lambda^*) = f(x^*)$