Optimization over a Convex Set

Lecture 9-1 - CMSE 382

Prof. Elizabeth Munch

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Fri, Mar 20, 2026

This Lecture

Topics:

- Stationarity
- Stationarity in convex problems
- Orthogonal projection revisited
- Gradient projection method

Announcements:

- Homework 4 due TODAY.

Section 1

Stationarity

Recall: Stationary point of a function in unconstrained optimization

Consider the **UN**constrained optimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

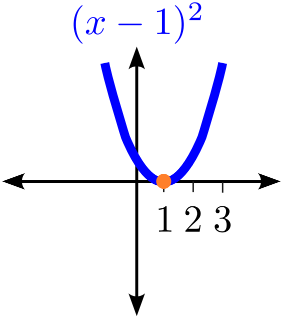
Definition (Stationary point of a function)

Let $f : U \to \mathbb{R}$ be a function defined on a set $U \subseteq \mathbb{R}^n$. Suppose that $\mathbf{x}^* \in \text{int}(U)$ and that f is differentiable over some neighborhood of $\mathbf{x}^*$. Then $\mathbf{x}^*$ is called a **stationary point of** f if $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

- It is a point where the gradient vanishes.

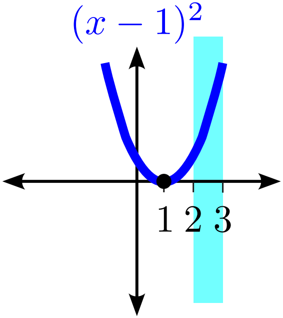
Stationary point of a function versus stationary point of a problem

$$\min_{x\in\mathbb{R}} (x-1)^2$$



$$x^* = 1$$

$$\min\{(x-1)^2 : x \in [2,3]\}$$



$$1 \notin [2,3]$$

Stationary point of a problem in constrained optimization

Consider the constrained optimization problem (P):

$$(P) \qquad \begin{aligned} &\text{minimize} && f(\mathbf{x}) \\ &\text{such that} && \mathbf{x} \in C \end{aligned}$$

Definition (Stationarity condition for a problem)

Let f be a continuously differentiable function over a closed convex set C. Then $\mathbf{x}^* \in C$ is called a stationary point of (P) if $\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0$ for any $\mathbf{x} \in C$.

- A point where there are no feasible descent directions.

Theorem (Stationarity as a necessary optimality condition)

Let f be a continuously differentiable function over a closed convex set C, and let $\mathbf{x}^*$ be a local minimum of (P). Then $\mathbf{x}^*$ is a stationary point of (P).

Equivalence of stationarity definitions when $C = \mathbb{R}^n$

Consider the optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{such that} \quad & \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

Stationary points for the problem satisfy:
$\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0$ for all $\mathbf{x}$

- Choose $\mathbf{x} = \mathbf{x}^* - \nabla f(\mathbf{x}^*)$
  - $\nabla f(\mathbf{x}^*)^\top (\mathbf{x}^* - \nabla f(\mathbf{x}^*) - \mathbf{x}^*)$
    $= -\nabla f(\mathbf{x}^*)^\top \nabla f(\mathbf{x}^*)$
    $= -\|\nabla f(\mathbf{x}^*)\|^2 \geq 0$
  - But $-\|\cdot\|^2 \leq 0$.
- So $\nabla f(\mathbf{x}^*) = \mathbf{0}$
- Stationarity definitions for a constrained minimization problem and an unconstrained problem coincide when the feasible region becomes $\mathbb{R}^n$.

Some special cases

Feasible set	Explicit Stationarity Condition
$C = \mathbb{R}^n$	$\nabla f(\mathbf{x}^*) = \mathbf{0}$
$C = \mathbb{R}^n_+$	$\begin{cases} \frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0, & x_i^* > 0 \\ \frac{\partial f}{\partial x_i}(\mathbf{x}^*) \geq 0, & x_i^* = 0 \end{cases}$
$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{e}^\top \mathbf{x} = 1\}$	$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) = \ldots = \frac{\partial f}{\partial x_n}(\mathbf{x}^*)$
$B[0,1]$	$\nabla f(\mathbf{x}^*) = \mathbf{0}$ or $\lVert \mathbf{x}^* \rVert = 1$ and $\exists \lambda \leq 0 : \nabla f(\mathbf{x}^*) = \lambda \mathbf{x}^*$

Stationary point of a convex problem in constrained optimization

Consider the constrained optimization problem (P):

$$\text{(P)} \quad \begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{such that} & \mathbf{x} \in C \end{array}$$

where C is convex.

Theorem (Stationarity as a necessary optimality condition)

Let f be a continuously differentiable function over a closed convex set C, and let $\mathbf{x}^*$ be a local minimum of (P). Then $\mathbf{x}^*$ is a stationary point of (P).

Theorem (Stationarity as necessary and sufficient condition for convex objective function)

Let f be a continuously differentiable convex function over a closed and convex set $C \subseteq \mathbb{R}^n$. Then $x^* \in C$ is a stationary point of (P) if and only if x^* is an optimal solution of (P).

Section 2

Gradient Projection Method

Recall: Orthogonal projection

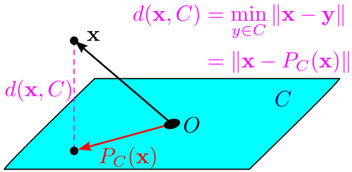
Definition (Recall: Orthogonal projection operator)

Given a nonempty closed convex set C, the orthogonal projection operator $P_C : \mathbb{R}^n \to C$ is defined by

$$P_C(\mathbf{x}) = \arg\min \|\mathbf{y} - \mathbf{x}\|^2 : \mathbf{y} \in C.$$

Theorem (Recall: first projection theorem)

Let C be a nonempty closed convex set. Then the problem $P_C(\mathbf{x}) = \arg\min \|\mathbf{y} - \mathbf{x}\|^2 : \mathbf{y} \in C$ has a unique optimal solution.



- Returns the vector in C that is closest to input vector $\mathbf{x}$.
- Is a convex optimization problem:

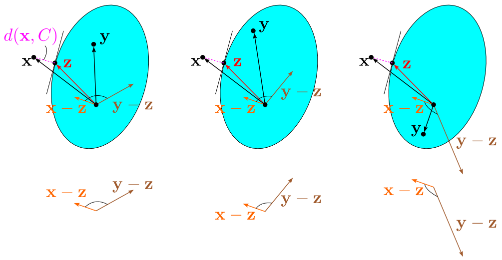
$$\begin{aligned} \min \quad & \|\mathbf{y} - \mathbf{x}\|^2 \\ \text{s.t.} \quad & \mathbf{y} \in C. \end{aligned}$$

Orthogonal projection: Second projection theorem

Theorem (Second projection theorem)

Let C be a closed convex set and let $\mathbf{x} \in \mathbb{R}^n$. Then $z = P_C(\mathbf{x})$ if and only if $z \in C$ and $(\mathbf{x} - \mathbf{z})^\top(\mathbf{y} - \mathbf{z}) \leq 0$ for any $y \in C$.

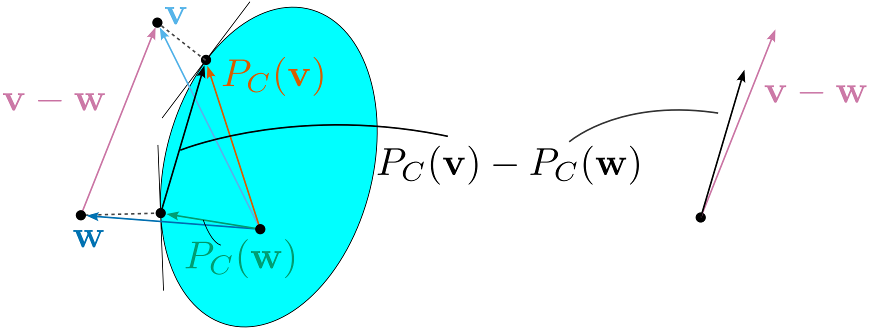
- The angle between $\mathbf{x} - P_C(\mathbf{x})$ and $\mathbf{y} - P_C(\mathbf{x})$ is greater than or equal to 90 degrees.



$\mathbf{a}^\top\mathbf{b} > 0$ $0 \leq \theta < \pi/2$	$\mathbf{a}^\top\mathbf{b} = 0$ $\theta = \pi/2$	$\mathbf{a}^\top\mathbf{b} < 0$ $\pi/2 < \theta \leq \pi$

$d(\mathbf{x}, C) = \|\mathbf{x} - \mathbf{z}\|$

Orthogonal projection



Theorem

Let C be a nonempty closed and convex set. Then

1. *for any* $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$: $(P_C(\mathbf{v}) - P_C(\mathbf{w}))^\top(\mathbf{v} - \mathbf{w}) \geq \|P_C(\mathbf{v}) - P_C(\mathbf{w})\|^2$.
2. *(non-expansiveness)*

$$\|P_C(\mathbf{v}) - P_C(\mathbf{w})\| \leq \|\mathbf{v} - \mathbf{w}\|$$

Representation of stationarity using the orthogonal projection operator

Theorem (Stationarity in terms of the orthogonal projection operator)

Let f be a continuously differentiable function defined on the closed and convex set C, and let $s > 0$. Then $\mathbf{x}^* \in C$ is a stationary point of the problem

$$(P) \quad \begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C. \end{array}$$

if and only if

$$\mathbf{x}^* = P_C(\mathbf{x}^* - s\nabla f(\mathbf{x}^*)).$$

- Can we use inspiration from this theorem for finding stationary points of optimization probelms over convex sets
  - Yes, the gradient projection method

Gradient projection algorithm

Gradient projection algorithm

Input: tolerance parameter $\varepsilon > 0$.
Initialization: Pick $x_0 \in C$ arbitrarily.
For any $k = 0, 1, 2, \ldots$ **do**:

1. Pick a stepsize t_k by a line search procedure.
   - For example, using fixed step size, exact line search, or backtracking.
2. Set $\mathbf{x}_{k+1} = P_C(\mathbf{x}_k - t_k \nabla f(\mathbf{x}_k))$.
3. If $\|\mathbf{x}_k - \mathbf{x}_{k+1}\| \leq \varepsilon$, then STOP and $\mathbf{x}_{k+1}$ is the output.

- In the unconstrained case, this is the same as gradient descent.
- There are convergence results.

Groups - Round 4

Group 1
Michal, Joseph, Saitej, Dev

Group 2
Kyle, Dori, Shanze, Jack

Group 3
Noah, Daniel, Lora, Scott

Group 4
Lowell, Tianjian, Aidan, Anthony

Group 5
Abigail, Breena, Arjun, Luis

Group 6
Purvi, Atticus, Andrew, Vinod

Group 7
Yousif, Jay, Arya, Morgan

Group 8
Jonid, Jake, Dominic, Maye

Group 9
Alice, K M Tausif, Monirul Amin, Ha

Group 10
Jamie, Zheng, Aaron, Long

Group 11
Lauryn, Karen, Sanskaar, Braedon

Group 12
Sai, Brandon, Igor, Quang Minh