CE 272
Traffic Network Equilibrium

Lecture 2

Review of Convex Optimization - Part I

Previously on Traffic Network Equilibrium...

Nash Equilibrium (1951)

At equilibrium, no player has an incentive to deviate.

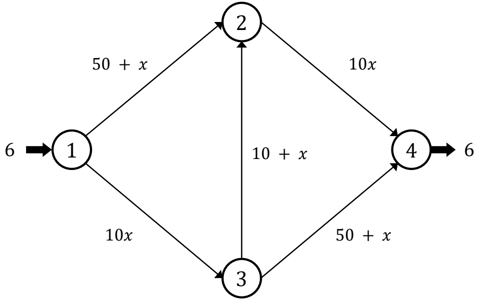
- **Who are the agents/players of the game?**
  Travelers
- **What are the actions available to each player?**
  Paths between the origin-destination (OD) pair
- **What are the (dis)utilities of each player for a particular outcome?**
  Path travel times

Wardrop Equilibrium (1952)

All used paths have equal and minimal travel time.

Previously on Traffic Network Equilibrium...

If you build it, they will come. But...



Building new road(s) may lead to more congestion. Equivalently, shutting down a road(s) may improve traffic!

Previously on Traffic Network Equilibrium...

Any optimization program can be written in the following standard form.

$$
\begin{aligned}
&\min_{\mathbf{x}} f(\mathbf{x}) && \\
\text{s.t. } & g_i(\mathbf{x}) \leq 0 && \forall\, i = 1, 2, \ldots, l \\
& h_i(\mathbf{x}) = 0 && \forall\, i = 1, 2, \ldots, m \\
& \mathbf{x} \in X &&
\end{aligned}
$$

The functions g and h define the inequality and equality constraints respectively. The set X is used to represent additional constraints (e.g., integrality), which we won't have in this course. Instead, we use X to denote implicit constraints.

The set of decision variables that satisfy all the constraints is called the **feasible region**.

Maximization problems can be converted into the standard form by simply changing the sign of the objective function.

Lecture Outline

Lecture Outline

Convex Sets

Convex Sets

Note on Notation

Throughout the course we will *try* to use

- Boldfaced lower case letters to denote vectors (e.g., $\mathbf{x}, \mathbf{y}$)
- Boldfaced upper case letters to denote matrices (e.g., $\mathbf{A}, \mathbf{B}$)
- Blackboard bold typeface to denote standard sets such as reals, integers (e.g., $\mathbb{R}, \mathbb{Z}$)
- Upper case (e.g., X, Y) and calligraphic letters (e.g., $\mathcal{M}, \mathcal{N}$) to denote sets and functions

Also, all vectors are assumed to be column vectors.

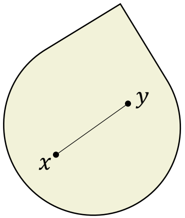
Convex Sets

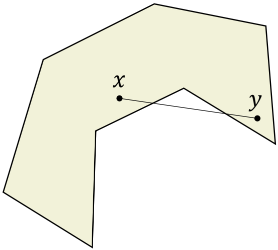
Definition

Definition (Convex Set)

A set X is convex iff the convex combination of any two points in the set also belongs to the set. Mathematically,

$$X \subseteq \mathbb{R}^n \text{ is convex} \Leftrightarrow \forall\, \mathbf{x}, \mathbf{y} \in X \text{ and } \forall\, \lambda \in [0,1], \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in X$$





Convex Sets

Examples

Which of the following sets are convex?

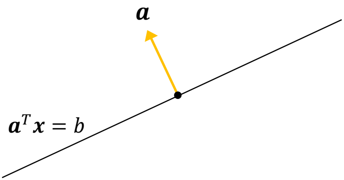
1. Empty Set $\varnothing$
2. Euclidean Ball $B(\mathbf{x}_0, \epsilon) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| \leq \epsilon\}$

Convex Sets

Examples

Definition (Hyperplane)

Sets of the form $\{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} = b\}$, where $\mathbf{a} \in \mathbb{R}^n, \mathbf{a} \neq \mathbf{0}, b \in \mathbb{R}$ are called hyperplanes.

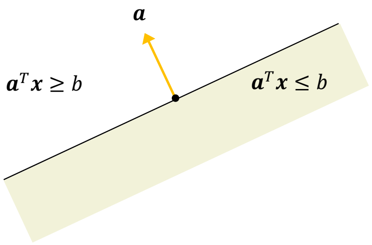


Convex Sets

Examples

Definition (Halfspace)

Sets of the form $\{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} \leq b\}$, where $\mathbf{a} \in \mathbb{R}^n, \mathbf{a} \neq \mathbf{0}, b \in \mathbb{R}$ are called halfspaces.

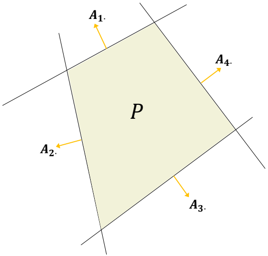


Convex Sets

Examples

Definition (Polyhedron)

A polyhedron is a set of the form $P = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ where $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m$.

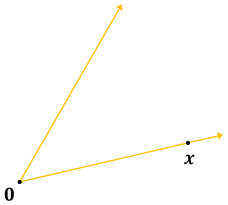


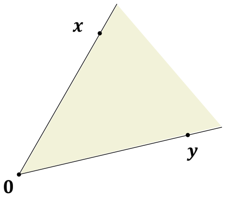
Definition (Polytope)

A bounded polyhedron is called a polytope.

Definition (Cone)

A set C is called a cone if for every $\mathbf{x} \in C$ and $\lambda \geq 0$, $\lambda \mathbf{x} \in C$.





Definition (Convex Cone)

A set C is called a convex cone if it is convex and a cone, i.e., $\forall\, \mathbf{x}, \mathbf{y} \in C$ and $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 \mathbf{x} + \lambda_2 \mathbf{y} \in C$.

Convex Functions

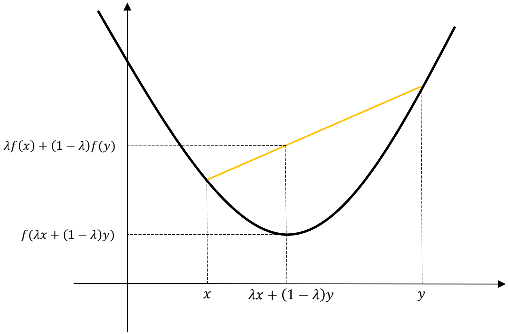
Convex Functions

Definition

Definition (Convex Function)

A function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex if $\forall \mathbf{x}, \mathbf{y} \in X, \lambda \in [0, 1]$,

$$f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$



Convex Functions

Examples

Which of the following functions are convex?

1. $12x$
2. x^3
3. x^4
4. $|x|$
5. e^x
6. $\sin x$
7. $\log x$
8. $f(\mathbf{x}) = \max\{x_1, x_2, \ldots, x_n\}$

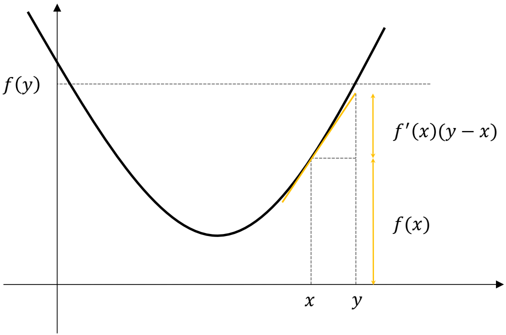
Convex Functions

Differentiable Functions

Definition (Convex Function)

A **differentiable** function of one variable $f : X \subseteq \mathbb{R} \to \mathbb{R}$ is convex iff

$$f(y) \geq f(x) + f'(x)(y - x) \,\forall\, x, y \in X$$



Convex Functions

Differentiable Functions

For functions of more than one variable, the derivative can be replaced with the gradient vector.

Definition (Convex Function)

A **differentiable** function of multiple variables $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is convex iff

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \, \forall \, \mathbf{x}, \mathbf{y} \in X$$

where

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Convex Functions

Differentiable Functions - Visualizing the Gradient

For a function of one variable, the gradient at a point in the domain is simply the slope of the tangent at the corresponding function value.

What does the gradient look like for functions of more than one variable?

Definition (Level Sets)

Let $c \in \mathbb{R}$. The level set of a function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is defined as

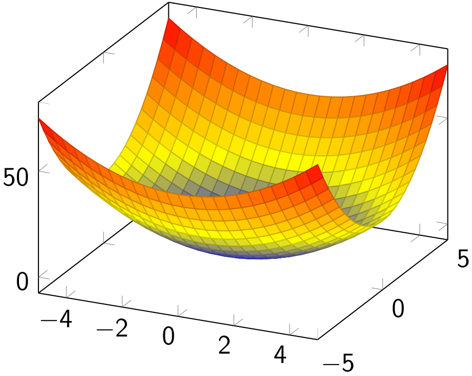
$$L_f(c) = \{\mathbf{x} \in X \mid f(\mathbf{x}) = c\}$$

When f is a function of two variables, the level sets are also referred to as **contour lines** or **level curves** and when f is a function of three variables they are called **level surfaces**.

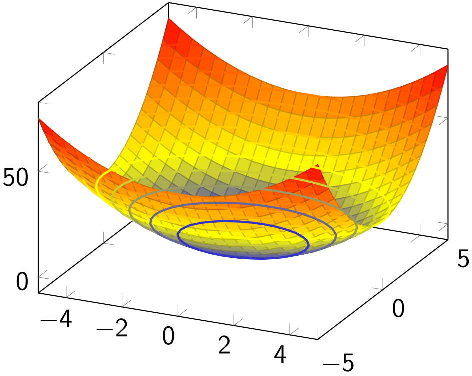
Convex Functions

Differentiable Functions - Visualizing the Gradient

Consider the function $f(\mathbf{x}) = x_1^2 + 2x_2^2$



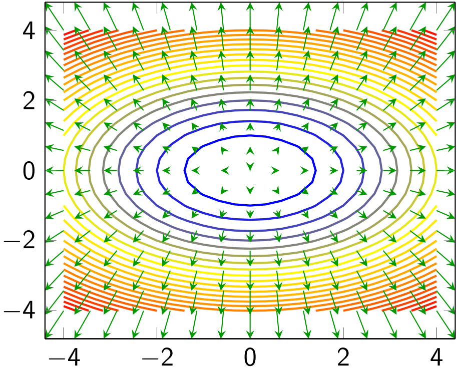
Consider the function $f(\mathbf{x}) = x_1^2 + 2x_2^2$. The level sets are ellipses $x_1^2 + 2x_2^2 = c$.



Convex Functions

Differentiable Functions - Visualizing the Gradient

The following plot shows the level sets and the gradient $[2x_1 \;\; 4x_2]^T$.



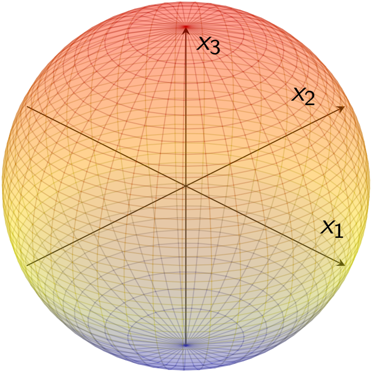
The gradient vector is 'orthogonal to the level sets'*

*A formal proof is a little involved but is not difficult and makes use of the Implicit Function Theorem.

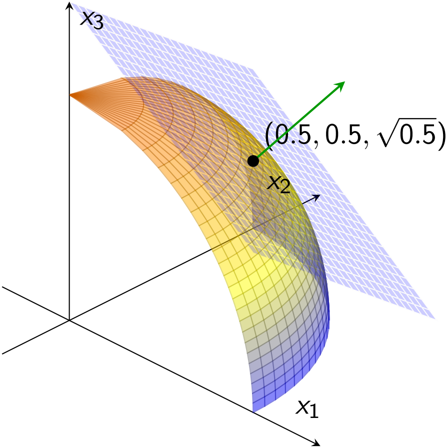
Convex Functions

Differentiable Functions - Visualizing the Gradient

Consider the function of three variables $f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$. (It's impossible to visualize this function as we can't create a 4D plot.) The level sets however can be drawn as they are spheres $x_1^2 + x_2^2 + x_3^2 = c$.



Consider the level set $x_1^2 + x_2^2 + x_3^2 = 1$. The gradient of the function is $[2x_1 \ 2x_2 \ 2x_3]^T$ and its value at $(0.5, 0.5, \sqrt{0.5})$ is $(1, 1, 2\sqrt{0.5})$, which is **normal to the tangent plane** at $(0.5, 0.5, \sqrt{0.5})$.

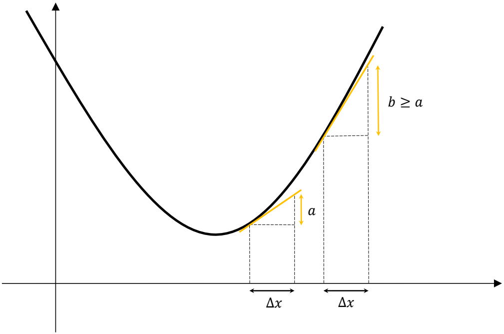


Convex Functions

Twice-Differentiable Functions

Definition (Convex Function)

A **twice-differentiable** function of one variable $f : X \subseteq \mathbb{R} \rightarrow \mathbb{R}$ is convex iff $f''(x) \geq 0 \, \forall \, x \in X$.



Convex Functions

Twice-Differentiable Functions

In higher dimensions, we make use of the Hessian matrix instead of the second derivative. But first, we need some additional definitions.

Definition (Positive Definite Matrix)

A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is called a positive definite matrix and is denoted as $\mathbf{A} \succ 0$ if $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$ for all non-zero $\mathbf{x} \in \mathbb{R}^n$.

Definition (Positive Semidefinite Matrix)

A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is called a positive semidefinite matrix and is denoted as $\mathbf{A} \succeq 0$ if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all non-zero $\mathbf{x} \in \mathbb{R}^n$.

One can similarly define *negative definite* and *negative semidefinite* matrices.

Definition (Convex Function)

A **twice-differentiable** function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex iff $\nabla^2 f(\mathbf{x}) \succeq 0 \, \forall \, \mathbf{x} \in X$.

where

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Convex Functions

Twice-Differentiable Functions

Verifying if the Hessian is positive semidefinite is challenging because we need to check if $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$ for all possible non-zero vectors $\mathbf{x}$.

Alternately, $\mathbf{A}$ with real eigenvalues is positive definite (semidefinite) iff all of its eigen values are positive (non-negative).

However, if the Hessian is a diagonal matrix the following proposition can be used.

Proposition (Positive Definite Matrix)

A diagonal matrix is positive definite (semidefinite) iff all diagonal elements are strictly positive (non-negative).

Convex Functions

Twice-Differentiable Functions

Which of the following functions are convex?

1. $x_1^2 + x_2^4 + x_3^2$
2. $x_1^2 + x_2^3 + x_2$

Convex Functions

Strict Convexity

Definition (Strict Convexity)

A function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is strictly convex if $\forall \mathbf{x}, \mathbf{y} \in X, \lambda \in [0, 1]$, $f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) < \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$

Definition (Strict Convexity)

A **differentiable** one-dimensional function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is strictly convex iff $f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \forall \mathbf{x}, \mathbf{y} \in X$

Definition (Strict Convexity)

A **twice-differentiable** function $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ is strictly convex iff $\nabla^2 f(\mathbf{x}) \succ 0 \forall \mathbf{x} \in X$.

Lecture Outline

Unconstrained Optimization

Unconstrained Optimization

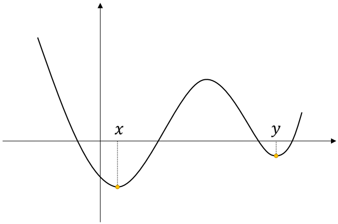
Local and Global Minima

Definition (Global Minimum)

$\mathbf{x}^* \in X$ is a global minimum of $f : X \to \mathbb{R}$ iff $f(\mathbf{x}^*) \leq f(\mathbf{x}) \, \forall \, \mathbf{x} \in X$.

Definition (Local Minimum)

$\mathbf{x}^* \in X$ is a local minimum of $f : X \to \mathbb{R}$ if $\exists \, \epsilon > 0$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x}) \, \forall \, \mathbf{x} \in X \cap B(\mathbf{x}^*, \epsilon)$.



Both x and y are local minima but only point x is the global minimum of f.

Unconstrained Optimization

Optimality Conditions

Before learning to minimize a function, we'd first like to *characterize* the optimal solutions. We will look at methods to find these solutions much later.

In other words, we seek answers to the following questions.

- **Are there conditions that the optimal solutions satisfy?**
- **If I gave you a feasible solution, can you tell me if it is optimal? (Certificate of Optimality)**

These type of results are also called **optimality conditions**. Note that the two questions are not synonymous. The conditions satisfied by optimal solutions may be also be satisfied by non-optimal solutions.

Unconstrained Optimization

Optimality Conditions

Let's answer these questions for unconstrained problems. (We treat problems with implicit constraints as unconstrained programs.)

Proposition (Necessary Conditions)

$\mathbf{x}^*$ *is a local minimum of a differentiable function* $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$
$\Rightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$

The converse is not true! (Why?) Thus, for $\mathbf{x}^*$ to be optimal it is not sufficient if $\nabla f(\mathbf{x}^*) = \mathbf{0}$. To derive sufficient conditions, we need convextiy.

Proposition (Necessary and Sufficient Conditions)

$\mathbf{x}^*$ *is a global minimum of a differentiable convex function*
$f : X \subseteq \mathbb{R}^n \to \mathbb{R} \Leftrightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$

These are also referred to as **first-order optimality conditions** since they use the first derivatives.

Unconstrained Optimization

Remarks on Existence of Optimal Solutions

A function may not have a minimum or may be unbounded. For example,

- e^x does not have a minimum (note that infimum exits)
- x is unbounded

However, if the problem is constrained, we could come up with conditions that guarantee the existence of an optima. E.g.,

Proposition

A continuous function with a compact domain (closed and bounded) always has a minimum.

Lecture Outline

Constrained Optimization

Constrained Optimization

Optimality Conditions

Recall our motivating questions:

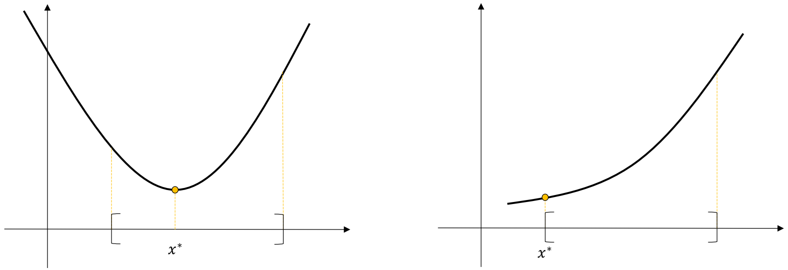
- **Are there conditions that the optimal solutions satisfy?**
- **If I gave you a feasible solution, can you tell me if it is optimal?**

Will our first-order conditions that were developed for the unconstrained case work when there are constraints?

Constrained Optimization

Optimality Conditions

Consider the following examples



If the minimum occurs at an interior point, it appears that we could use the conditions developed earlier.

But if a corner point is a minima, the gradient need not be zero!

For constrained problems, there are two ways of characterizing the optimal solutions.

- The first one is simple, easy to derive and understand, but is not very helpful.
- The second approach on the other hand requires a new concept called duality, is more involved, and more insightful!

At the end of the next lecture, we will also see how these two are related.

Constrained Optimization

Standard Form

Recall our standard form of an optimization problem

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x}) && \\ \text{s.t. } & g_i(\mathbf{x}) \le 0 && \forall\, i = 1, 2, \ldots, l \\ & h_i(\mathbf{x}) = 0 && \forall\, i = 1, 2, \ldots, m \end{aligned}$$

Suppose the above problem is a convex program and, as before, assume that all the functions are differentiable.

Let $\mathcal{X}$ denote the set of $\mathbf{x}$ that satisfy the inequality and equality and other implicit constraints.

Constrained Optimization

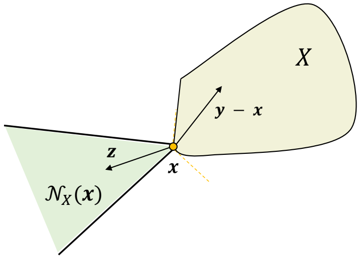
Normal Cone

Definition (Normal Cone)

Let $X \subseteq \mathbb{R}^n$, the *normal cone* of X at $\mathbf{x}$ is defined as

$$\mathcal{N}_X(\mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z}^T(\mathbf{y} - \mathbf{x}) \leq 0, \ \forall \ \mathbf{y} \in X\}$$

For the purpose of the following illustration, assume $\mathbf{x}$ is the origin.

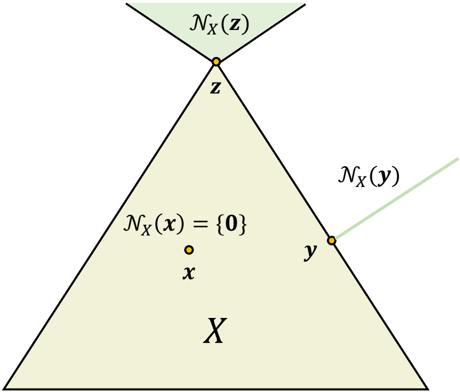


Vectors belonging to this set are called *normal vectors* to X at $\mathbf{x}$. They make a non-acute angle with $(\mathbf{y} - \mathbf{x})$ for all $\mathbf{y} \in X$.

What is the normal cone at an interior point?

Constrained Optimization

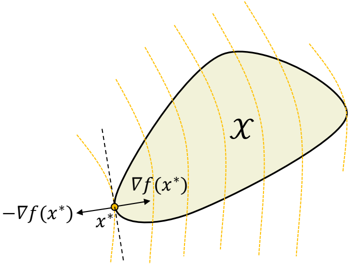
Normal Cone



Proposition (Necessary and Sufficient Conditions)

$\mathbf{x}^*$ *is an optimal to the convex program* $\min f(\mathbf{x})$ *s.t.* $\mathbf{x} \in \mathcal{X}$ *iff* $-\nabla f(\mathbf{x}^*) \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$

Consider a function of two variables with feasible region $\mathcal{X}$ and level sets as shown below. Recall that the gradient is orthogonal to the level curves.



For interior optima, the normal cone is $\{\mathbf{0}\}$, hence $\nabla f(\mathbf{x}^*) = \mathbf{0}$!

Proposition (Necessary and Sufficient Conditions)

$\mathbf{x}^*$ *is an optimal to the convex program* $\min f(\mathbf{x})$ *s.t.* $\mathbf{x} \in \mathcal{X}$ *iff* $-\nabla f(\mathbf{x}^*) \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$

Proof.

($\Leftarrow$) Note that since f is convex, the following is true,

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \, \forall \, \mathbf{y} \in \mathcal{X}$$

If $-\nabla f(\mathbf{x}^*) \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$,

$$\nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \geq 0 \, \forall \, \mathbf{y} \in \mathcal{X}$$

Therefore, $f(\mathbf{x}^*) \leq f(\mathbf{y}) \, \forall \, \mathbf{y} \in \mathcal{X}$ and hence $\mathbf{x}^*$ is optimal.

($\Rightarrow$) Exercise. $\square$

While the normal cone version of the optimality conditions $-\nabla f(\mathbf{x}^*) \in \mathcal{N}_{\mathcal{X}}(\mathbf{x}^*)$ is simple, it is not that useful.

Constructing the normal cone involves finding vectors which make a non-acute angle with every $(\mathbf{y} - \mathbf{x}^*)$, where $\mathbf{y} \in \mathcal{X}$!

Your Moment of Zen