Examples: Interior

Quiz: Find the interior of the following sets as a subset of $\mathbb{R}$.

1. $\{x \mid -1 \leq x \leq 1\}$
2. $\{1/n \mid n \in \mathbb{N}\}$

Quiz: Find the closure of the following sets as a subset of $\mathbb{R}$.

1. $\{x \mid 1 < x < 2\}$
2. $\{1, 2, 3\}$

$[1, 2]$

$1 \in B_\epsilon(1)$

$B_\epsilon(1) \cap \{1,2,3\} = \{1\}$

$B_\epsilon(1) \cap \{1,2,3\}^c$

$S \cup B_{ry}$

$B_\epsilon(1) \cap \{1,2,3\} \neq \emptyset$

$\{1,2,3\} \cup$

pts in this set s.t. any B_ϵ s.t. $B_\epsilon \cap \{1,2,3\} \neq \emptyset$ & $B_\epsilon \cap \{1,2,3\}^c \neq \emptyset$

Examples: Interior

Quiz: Find the interior of the following sets as a subset of $\mathbb{R}$.

1. $\{x \mid -1 \leq x \leq 1\}$
2. $\{1/n \mid n \in \mathbb{N}\}$

Quiz: Find the closure of the following sets as a subset of $\mathbb{R}$.

1. $\{x \mid 1 < x < 2\}$
2. $\{1, 2, 3\}$

$B_\epsilon(x) \cap \mathbb{Q} \neq \phi$

$B_\epsilon(x) \cap \mathbb{Q}^c \neq \phi$

$\epsilon > 0$

Quiz: Find the boundary of the following sets as a subset of $\mathbb{R}$.

1. $\{x \mid -1 < x < 1\}$ $\quad \{-1, 1\}$
2. $\mathbb{Q}$

Boundary$(\mathbb{Q}) = \mathbb{R}$

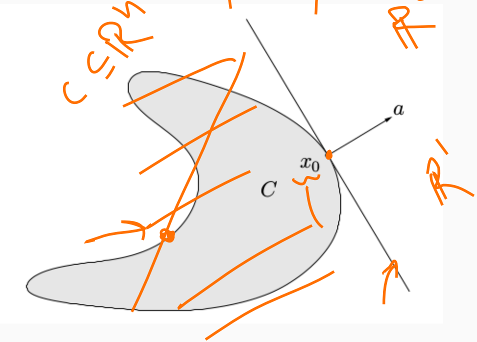
Supporting Hyperplanes

Supporting Hyperplane: Suppose $C \subseteq \mathbb{R}^n$, and x_0 is a point in its boundary **bd** C, i.e.,

$$x_o \in \textbf{bd } C = \textbf{cl } C \setminus \textbf{int } C$$

If $a \neq 0$ satisfies $a^T x \leq a^T x_0$ for all $x \in C$, then the hyperplane $\{x \mid a^T x = a^T x_0\}$ is called the supporting hyperplane to C at the point x_0. This is same as saying that the point x_0 is separated by the hyperplane $\{x \mid a^T x = a^T x_0\}$. The geometric interpretation is that the hyper-plane $\{x \mid a^T x = a^T x_0\}$ is tangent to C at x_0, and the halfspace $\{x \mid a^T x \leq a^T x_0\}$ contains C.

Figure 2: Supporting Hyperplane



History of Supporting Hyperplane

- Hermann Minkowski introduced separating and supporting hyperplane theorem
- He also extensively used convex sets and convex functions
  - Convex sets were first used by Archimedes, a Greek mathematician
- Photo on the right during Mathematics prize by French academy of sciences
- Minkowski used so called convex bodies concept to prove many results in number theory
  - See "Geometry of numbers" by H. Minkowski

Figure 3: Hermann Minkowski, 1883

Die nirgends concaven Flächen.

1. Es sei ein Punkt e und eine Punktmenge F mit folgenden Eigenschaften gegeben:

(A.) In jeder Richtung von e aus soll mindestens ein Punkt von F liegen (es ist ein Punkt mit bestimmten endlichen Coordinaten und verschieden von e gemeint, denn zu einer Richtung bedarf es zweier verschiedener Punkte, vgl. 2).

18.

Die überall convexen Flächen.

Eine nirgends concave Fläche soll als überall convex bezeichnet werden, wenn jede Stützebene an die Fläche mit derselben nur einen Punkt gemein hat. Diese Bedingung lässt sich noch anders ausdrücken.

Es sei F eine nirgends concave Fläche um einen Punkt e. Existirt eine Stützebene an F, welche zwei verschiedene Punkte von F ent-

Figure 4: Convex and Strictly Convex from the article: Geometry of numbers

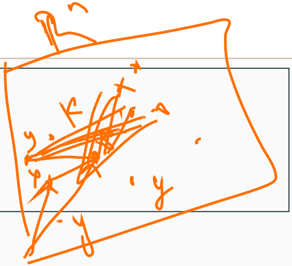
Dual Cones and Generalized Inequalities

$x^T y = y^T x$

Dual Cone: Let K be a cone. The set

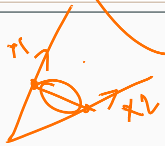
$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$$

is called the dual cone of K.



Dual Cones and Generalized Inequalities

$K = \{\theta x_1\} \cup \{\theta x_2\}$



Dual Cone: Let K be a cone. The set

$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$$

is called the dual cone of K.

- K^* is a cone, and is always convex, even when the original cone K is not
  - Proof on chalkboard

K^*

$x_1 \in K^*$, $x_2 \in K^*$, $x \in K$

$\theta_1 x_1 + \theta_2 x_2 \in K^*$

$\theta_1 x_1^T x \geq 0 \quad \forall x \in K$

$\theta_2 x_2^T x \geq 0 \quad \forall x \in K$

$\theta_1 x_1^T x + \theta_2 x_2^T x \geq 0 \quad \forall x \in K$

$(\theta_1 x_1 + \theta_2 x_2)^T x \geq 0 \quad \forall x$

$\theta_1 x_1 + \theta_2 x_2 \in K^*$

min $f(x)$
s.t. $x \in K$

dual of $f(y)$

K^*

convex cone

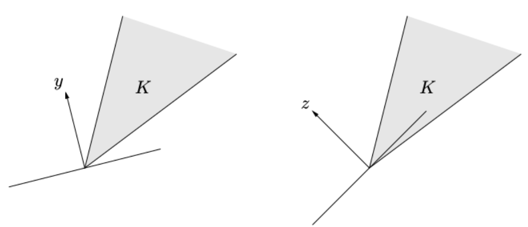
Dual Cones and Generalized Inequalities

Dual Cone: Let K be a cone. The set

$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$$

is called the dual cone of K.

- K^* is a cone, and is always convex, even when the original cone K is not
  - Proof on chalkboard



Dual Cones and Generalized Inequalities

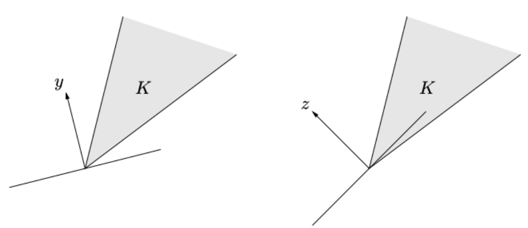
Dual Cone: Let K be a cone. The set

$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$$

is called the dual cone of K.

- K^* is a cone, and is always convex, even when the original cone K is not
  - Proof on chalkboard

$u^T v \leq 0$



- $x^T y \geq 0 \iff y^T x \geq 0 \iff y^T x \geq y^T 0 \iff (-y)^T x \leq (-y)^T 0$

Dual Cones and Generalized Inequalities

Dual Cone: Let K be a cone. The set

$$K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$$

is called the dual cone of K.

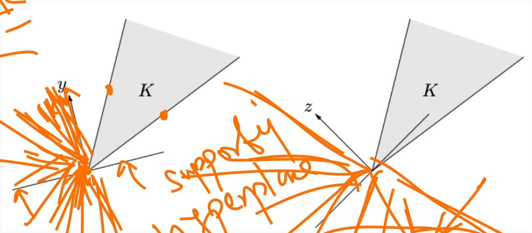
$x \in K$

$\theta x \in K$

$\theta > 0$

$\theta = 0$

- K^* is a cone, and is always convex, even when the original cone K is not
  - Proof on chalkboard



supporting hyperplane

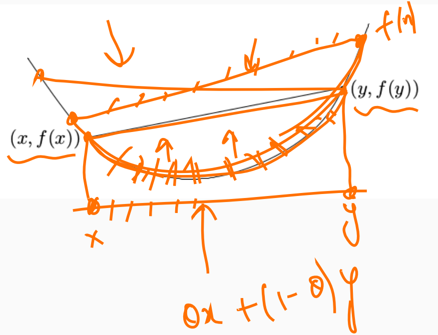
$a^T x \leq 0$

- $x^T y \geq 0 \iff y^T x \geq 0 \iff y^T x \geq y^T 0 \iff (-y)^T x \leq (-y)^T 0$
- Dual cone contains set of all y such that $-y$ is the normal of the hyperplane that supports K

Convex Functions, Concave Functions

> **Convex function:** A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if **dom** f is a convex set and if for all $x, y \in$ **dom** f, and θ with $0 \leq \theta \leq 1$, we have
>
> $$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

Figure 5: Convex function

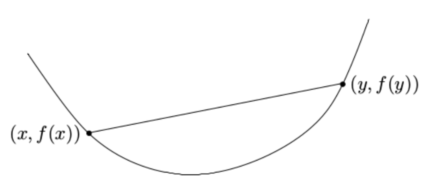


Convex Functions, Concave Functions

Convex function: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if **dom** f is a convex set and if for all $x, y \in$ **dom** f, and θ with $0 \le \theta \le 1$, we have

$$f(\theta x + (1 - \theta y)) \le \theta f(x) + (1 - \theta) f(y)$$

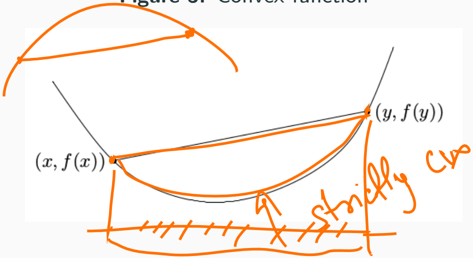
Figure 5: Convex function



Geometrically: The line segment joining points $(x, f(x))$ and $(y, f(y))$ lies above the graph of f.

Convex Functions, Concave Functions

Figure 5: Convex function

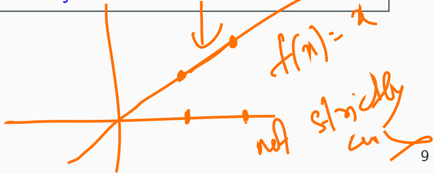


Convex function: A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if **dom** f is a convex set and if for all $x, y \in$ **dom** f, and θ with $0 \leq \theta \leq 1$, we have

$$f(\theta x + (1 - \theta y)) \leq \theta f(x) + (1 - \theta) f(y)$$

Geometrically: The line segment joining points $(x, f(x))$ and $(y, f(y))$ lies above the graph of f.

Strictly convex: A function is strictly convex if strict inequality holds whenever $x \neq y$ and $0 < \theta < 1$. We say f is concave if $-f$ is convex, and strictly concave if $-f$ is strictly convex



Examples of Convex Functions

$g'(x) > 0$

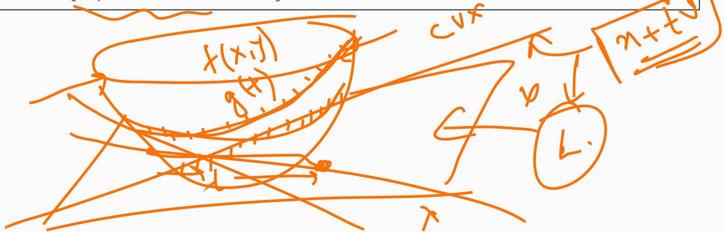
Quiz: Recall that an affine function is of the form $f(x) = Ax - b$. Is this a convex function? What about linear function $f(x) = Ax$?

Proof: on chalkboard

Fact: A function is convex if and only if it is convex when restricted to any line that intersects its domain.

In other words, a function is convex if and only if for all $x \in \textbf{dom } f$ and all v, the function $g(t) = f(x + tv)$ is convex on its domain $\{t \mid x + tv \in \textbf{dom } f\}$

Proof: On chalkboard.



Scratch Space

→ Since f is convex.

Claim: $g(t)$ is convex

Let $t_1, t_2 \in \mathbb{R}$.

$\theta g(t_1) + (1-\theta) g(t_2)$

$= \theta f(\underbrace{x + t_1 y}) + (1-\theta) f(\underbrace{x + t_2 y})$

$\geq f\big(\theta(x+t_1 y) + (1-\theta)(x+t_2 y)\big)$ (because f is cvx)

$= f(x + \underbrace{(\theta t_1 + (1-\theta)t_2)}y) = g(\theta t_1 + (1-\theta)t_2)$

$\Rightarrow g$ is convex

Scratch Space

Scratch Space

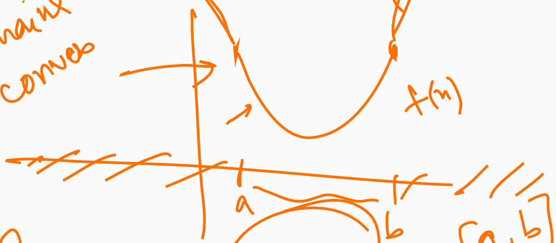
Extended Value Extensions

Extended value function: If f is convex we define its extended value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \infty$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \textbf{dom } f \\ \infty & x \notin \textbf{dom } f. \end{cases}$$

remains convex

$f : \mathbb{R}^n \to \mathbb{R}$



cvx

$\tilde{f} : \mathbb{R}^n \to \mathbb{R}$

cvx

dom f = $[a, b]$

$f : [a, b] \to \mathbb{R}$

$f : \mathbb{R} \to \mathbb{R}$

Extended Value Extensions

Extended value function: If f is convex we define its extended value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \infty$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \textbf{dom } f \\ \infty & x \notin \textbf{dom } f. \end{cases}$$

- $\tilde{f}$ is defined on all $\mathbb{R}^n$ and takes values in $\mathbb{R} \cup \infty$

Extended Value Extensions

Extended value function: If f is convex we define its extended value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \infty$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \mathbf{dom}\ f \\ \infty & x \notin \mathbf{dom}\ f. \end{cases}$$

- $\tilde{f}$ is defined on all $\mathbb{R}^n$ and takes values in $\mathbb{R} \cup \infty$
- domain of f can be recovered from $\tilde{f}$ as $\mathbf{dom}\ f = \{x \mid \tilde{f}(x) < \infty\}$

Extended Value Extensions

Extended value function: If f is convex we define its extended value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \infty$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \textbf{dom } f \\ \infty & x \notin \textbf{dom } f. \end{cases}$$

- $\tilde{f}$ is defined on all $\mathbb{R}^n$ and takes values in $\mathbb{R} \cup \infty$
- domain of f can be recovered from $\tilde{f}$ as $\textbf{dom } f = \{x \mid \tilde{f}(x) < \infty\}$
- It simplifies. No need to explicitely describe the domain

Extended Value Extensions

Extended value function: If f is convex we define its extended value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \infty$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \textbf{dom } f \\ \infty & x \notin \textbf{dom } f. \end{cases}$$

- $\tilde{f}$ is defined on all $\mathbb{R}^n$ and takes values in $\mathbb{R} \cup \infty$
- domain of f can be recovered from $\tilde{f}$ as $\textbf{dom } f = \{x \mid \tilde{f}(x) < \infty\}$
- It simplifies. No need to explicitely describe the domain
- Convexity of a function can now be defined without specifying domain:

$$\tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1 - \theta \tilde{f}(y))$$

Extended Value Extensions

Extended value function: If f is convex we define its extended value extension $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \infty$ by

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \textbf{dom } f \\ \infty & x \notin \textbf{dom } f. \end{cases}$$

- $\tilde{f}$ is defined on all $\mathbb{R}^n$ and takes values in $\mathbb{R} \cup \infty$
- domain of f can be recovered from $\tilde{f}$ as $\textbf{dom } f = \{x \mid \tilde{f}(x) < \infty\}$
- It simplifies. No need to explicitely describe the domain
- Convexity of a function can now be defined without specifying domain:

$$\tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1 - \theta \tilde{f}(y))$$

- If x, y are not in $\textbf{dom } f$, then right hand side is ∞

Indicator Function and Extended Values

Indicator Function: Let $C \subseteq \mathbb{R}^n$ be a convex set, and consider the function I_C with domain C and $I_C(x) = 0$ for all $x \in C$. In other words, the function is identically zero on the set C. Its extended value extension is given by

$$\tilde{I}_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

The convex function $\tilde{I}_C$ is called the indicator function of the set C

Indicator Function and Extended Values

Indicator Function: Let $C \subseteq \mathbb{R}^n$ be a convex set, and consider the function I_C with domain C and $I_C(x) = 0$ for all $x \in C$. In other words, the function is identically zero on the set C. Its extended value extension is given by

$$\tilde{I}_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

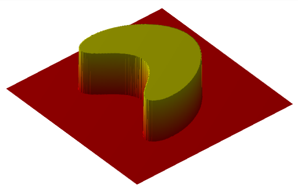
The convex function $\tilde{I}_C$ is called the indicator function of the set C

- The problem of minimizing a function f on the set C is the same as minimizing the function $f + \tilde{I}_C$ over all of $\mathbb{R}^n$
- Proof on chalkboard!

$\tilde{f} = f + \tilde{I}_c = f(x)$

$= \infty$

Figure 6: Indicator Function



Convexity of Indicator Function $\equiv$ Set is convex

Fact: A set is convex if and only if the indicator function is a convex function

Proof on chalkboard! $\Rightarrow$

C cvx set $\Rightarrow$ $\tilde{I}_C$ cvx fn

$x, y \in \mathbb{R}^n$

$$\theta \tilde{I}_C(x) + (1-\theta)\tilde{I}_C(y) \geq \tilde{I}_C\big(\underbrace{\theta x + (1-\theta) y}_{\in C}\big)$$

$\in C$ because C is cvx

Case-1: $x, y \in C$

$\theta \cdot 0 + (1-\theta)\cdot 0 = 0 \geq 0$ ✓

Case-2: $x \in C$, $y \notin C$

$\theta \cdot 0 + (1-\theta)\cdot \infty \geq \infty$ ✓

Case-3: $x \notin C$, $y \notin C$

$x \notin C$, $y \notin C$

$\infty \geq \infty$ ✓

Scratch Space

$\tilde{I}_C$ cvx $\Rightarrow$ C is cvx

$\tilde{I}_C(x_1) = \tilde{I}_C(x_2) = 0$

$x_1, x_2 \in C \Rightarrow$

$\theta x_1 + (1-\theta) x_2 \in C$

$\Rightarrow \theta x_1 + (1-\theta) x_2 \in C$

$\Rightarrow C$ '

Claim $\tilde{I}_C(\theta x_1 + (1-\theta) x_2) = 0$

$\theta \tilde{I}_C(x_1) + (1-\theta) \tilde{I}_C(x_2) \geq \tilde{I}_C(\theta x_1 + (1-\theta) x_2)$

$0 \qquad 0$

$\Rightarrow \tilde{I}_C(\theta x_1 + (1-\theta) x_2) \leq 0 \Rightarrow \tilde{I}_C(\quad) = 0$

First Order Condition for Convexity

First Order Condition: Suppose f is differentiable. Then f is convex **if and only if dom** f is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

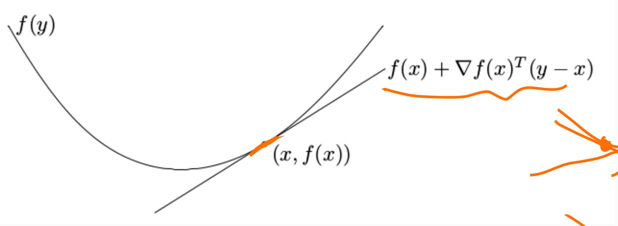
holds for all $x, y \in \textbf{dom } f$

First Order Condition for Convexity

First Order Condition: Suppose f is differentiable. Then f is convex **if and only if dom** f is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

holds for all $x, y \in$ **dom** f



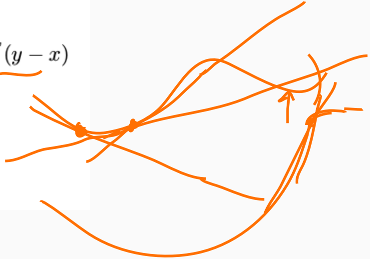


Figure 7: First order convexity

First Order Condition for Convexity

First Order Condition: Suppose f is differentiable. Then f is convex **if and only if dom** f is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

holds for all $x, y \in \textbf{dom } f$

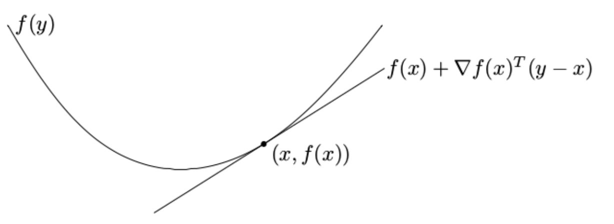


Figure 7: First order convexity

Let $f(y) = mx + c$ be the line passing through $(x, f(x))$. Then $c = f(x) - mx$, where $m = \nabla f(x)$. Hence, $f(y) = f(x) + \nabla f(x)(y - x)$ is the line that passes through $(x, f(x))$

Remarks on First Order Condition

- The Taylor expansion of infinitely differentiable function $f(x)$ at a point $x = a$

Remarks on First Order Condition

- The Taylor expansion of infinitely differentiable function $f(x)$ at a point $x = a$

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$
$$= f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots$$

$f^{(n)}$

Remarks on First Order Condition

- The Taylor expansion of infinitely differentiable function $f(x)$ at a point $x = a$

$$
\begin{aligned}
f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \\
&= f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots
\end{aligned}
$$

- Exercise: Expand the following functions using Taylor series
  - $f: \mathbb{R} \to \mathbb{R}$, Expand $f(x) = e^x$ at $\overset{a}{x} = 0$

$$f(0) + \frac{f'(0)}{1!}(x-0) + \frac{f''(a)}{2!}(x-0)^2$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Remarks on First Order Condition

- The Taylor expansion of infinitely differentiable function $f(x)$ at a point $x = a$

$$
\begin{aligned}
f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \\
&= f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \ldots
\end{aligned}
$$

- Exercise: Expand the following functions using Taylor series
  - $f : \mathbb{R} \to \mathbb{R}$, Expand $f(x) = e^x$ at $x = 0$
  - $f : \mathbb{R} \to \mathbb{R}$, Expand $f(x) = \sin x$ at $x = \pi$

Remarks on First Order Condition

- The Taylor expansion of infinitely differentiable function $f(x)$ at a point $x = a$

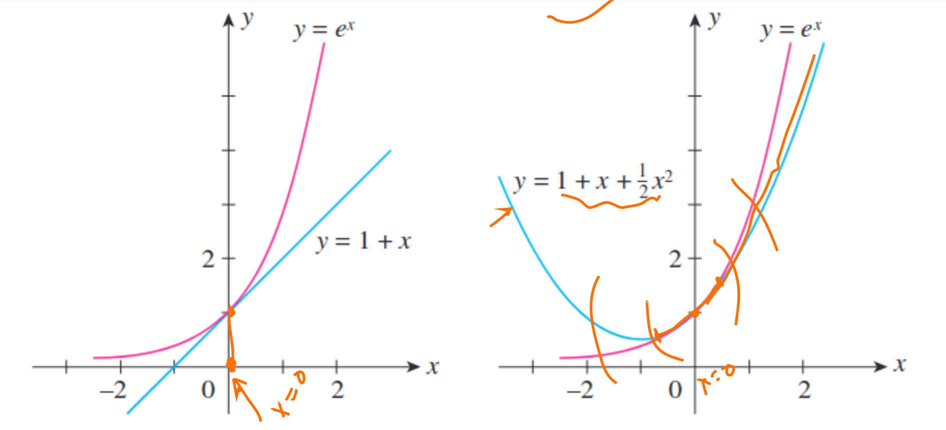
$$
\begin{aligned}
f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \\
&= f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots
\end{aligned}
$$

- Exercise: Expand the following functions using Taylor series
  - $f : \mathbb{R} \to \mathbb{R}$, Expand $f(x) = e^x$ at $x = 0$
  - $f : \mathbb{R} \to \mathbb{R}$, Expand $f(x) = \sin x$ at $x = \pi$
  - $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, Expand $f(x) = \dfrac{1}{1-x}$ at $x = -1$

$f : \mathbb{R} \to \mathbb{R}, \quad$ Expand $f(x) = e^x$ at $x = 0$

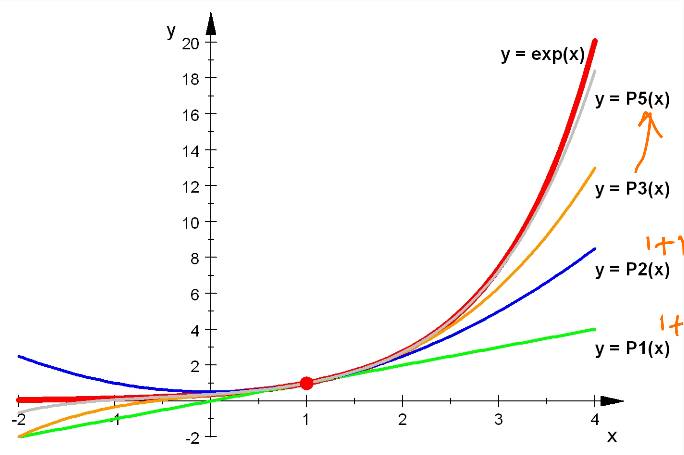
$f : \mathbb{R} \to \mathbb{R}$, Expand $f(x) = e^x$ at $x = 0$

$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$



$f : \mathbb{R} \to \mathbb{R}, \quad$ Expand $f(x) = e^x$ at $x = 1$

$f : \mathbb{R} \to \mathbb{R}, \quad$ Expand $f(x) = e^x$ at $x = 1$



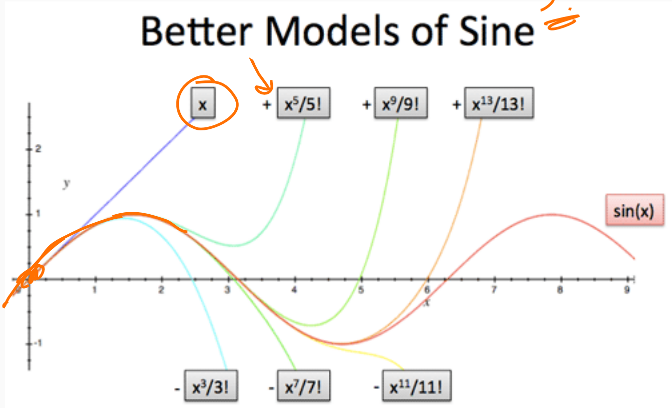
$1+x+\frac{x^2}{2}$

$1+x$

$$\boxed{f : \mathbb{R} \to \mathbb{R}, \quad \text{Expand } f(x) = \sin x \text{ at } x = 0}$$

$f : \mathbb{R} \to \mathbb{R}, \quad$ Expand $f(x) = \sin x$ at $x = 0$

$x + \frac{x^3}{3!} + \frac{x^5}{5!} \cdots$



Multivariate Taylor Series

$f : \mathbb{R} \to \mathbb{R}$.

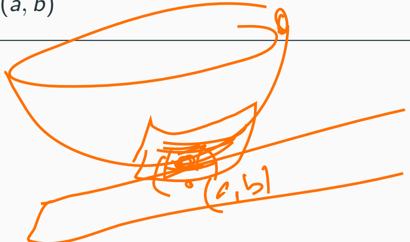
The Taylor expansion of function $f : \mathbb{R} \to \mathbb{R}$ near point (a, b) is given as follows:

$$f(x, y) = \underbrace{f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a)}_{\text{Plane}} + \frac{\partial f}{\partial x}(a, b)(x - b) + \frac{1}{2}\left(A(x - a)^2 + 2B(x - a)(y - b) + C(y - b)^2\right),$$

where

$$A = \frac{\partial^2 f}{\partial x^2}(a, b), \quad C = \frac{\partial^2 f}{\partial y^2}(a, b)$$

$$B = \frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$



Derivation of Multivariate Taylor Series...

Derivation of Multivariate Taylor Series...

Derivation of Multivariate Taylor Series...

Multivariate Taylor Series Using Gradient and Hessian...

The Taylor expansion of function $f : \mathbb{R} \to \mathbb{R}$ near point (a, b) is given as follows:

$$f(x, y) = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial x}(a, b)(x - b) + \frac{1}{2}(A(x - a)^2 + 2B(x - a)(y - b) + C(y - b)^2),$$

where

$$A = \frac{\partial^2 f}{\partial x^2}(a, b), \quad C = \frac{\partial^2 f}{\partial y^2}(a, b)$$

$$B = \frac{\partial^2 f}{\partial x \partial y}(a, b) = \frac{\partial^2 f}{\partial y \partial x}(a, b)$$

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

$$= f(a,b) + \nabla f(a,b)^T \begin{bmatrix} x-a \\ y-b \end{bmatrix} + [x-a, \; y-b]^T H \begin{bmatrix} x-a \\ y-b \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b) \end{bmatrix}$$

Taylor approximation

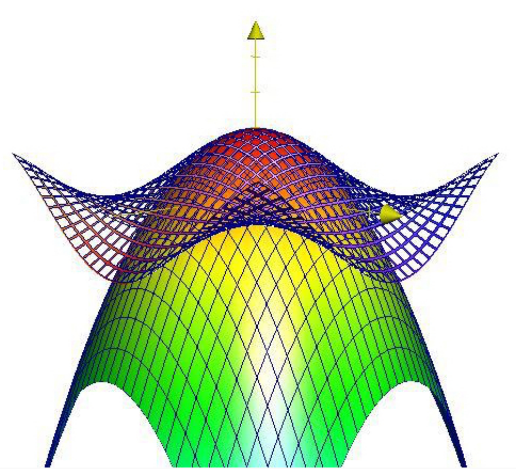


Figure 1: A second order Taylor approximation of the function

Remarks on First Order Condition

- The affine function of y given by $f(x) + \nabla f(x)^T(y - x)$ is a first order Taylor approximation of f near x.

$f(x)$ around $x = a$

$y = x$: $f(y) = f(x) + \nabla f(x)^T (y - x)$

- The affine function of y given by $f(x) + \nabla f(x)^T (y - x)$ is a first order Taylor approximation of f near x.
- The inequality

$f(a) + \nabla f(a)(x - a)$

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

$f(y)$ $y = a$

states that for a convex function the first order Taylor approximation is in fact a global underestimator of the function

Remarks on First Order Condition

- The affine function of y given by $f(x) + \nabla f(x)^T(y - x)$ is a first order Taylor approximation of f near x.
- The inequality

$$f(y) \geq f(x) + \nabla f(x)^T(y - x)$$

  states that for a convex function the first order Taylor approximation is in fact a global underestimator of the function
- Conversly, if the first order Taylor approximation of a function is always a global underestimator of the function, then the function is convex
- Quiz: What does the inequality above say when $\nabla f(x) = 0$?