# HW: Illustrating Computation of Directional Derivative

- As another example, let us find the rate of change of $f(x, y, z) = e^{xyz}$ at $p_0 = (1, 2, 3)$ in the direction from $p_1 = (1, 2, 3)$ to $p_2 = (-4, 6, -1)$.
- We first construct a unit vector from $p_1$ to $p_2$; $\mathbf{v} = \frac{1}{\sqrt{57}}[-5, 4, -4]$.
- The gradient of $f$ in general is $\nabla f = [yze^{xyz},\ xze^{xyz},\ xye^{xyz}] = e^{xyz}[yz,\ xz,\ xy]$.
- Evaluating the gradient at a specific point $p_0$, $\nabla f(1, 2, 3) = e^6\,[6, 3, 2]^T$. The directional derivative at $p_0$ in the direction $\mathbf{v}$ is $D_{\mathbf{u}} f(1, 2, 3) = e^6[6, 3, 2].\frac{1}{\sqrt{57}}[-5, 4, -4]^T = e^6\frac{-26}{\sqrt{57}}$.
- This directional derivative is negative, indicating that the function $f$ decreases at $p_0$ in the direction from $p_1$ to $p_2$.

# HW: Level Surface based Interpretation of Gradient: Examples

- Let $f(x_1, x, x_3) = x_1^2 x_2^3 x_3^4$ and consider the point $\mathbf{x}^0 = (1, 2, 1)$. We will find the equation of the tangent plane to the level surface through $\mathbf{x}^0$.
- The level surface through $\mathbf{x}^0$ is determined by setting $f$ equal to its value evaluated at $\mathbf{x}^0$; that is, the level surface will have the equation $x_1^2 x_2^3 x_3^4 = 1^2 2^3 1^4 = 8$.
- The gradient vector (normal to tangent plane) at $(1, 2, 1)$ is $\nabla f(x_1, x_2, x_3)\big|_{(1,2,1)} = [2x_1 x_2^3 x_3^4,\ 3x_1^2 x_2^2 x_3^4, 4x_1^2 x_2^3 x_3^3]^T\Big|_{(1,2,1)} = [16,\ 12,\ 32]^T$.
- The equation of the tangent plane at $\mathbf{x}^0$, given the normal vector $\nabla f(\mathbf{x}^0)$ can be easily written down: $\nabla f(\mathbf{x}^0)^T.[\mathbf{x} - \mathbf{x}^0] = 0$ which turns out to be $16(x_1 - 1) + 12(x_2 - 2) + 32(x_3 - 1) = 0$, a plane in $3D$.

# HW: Level Surface based Interpretation of Gradient: Examples

- Consider the function $f(x,y,z) = \frac{x}{y+z}$. The directional derivative of $f$ in the direction of the vector $\mathbf{v} = \frac{1}{\sqrt{14}}[1,\ 2,\ 3]$ at the point $x^0 = (4,1,1)$ is $\nabla^T f\Big|_{(4,1,1)} \cdot \frac{1}{\sqrt{14}}[1,\ 2,\ 3]^T = \left[\frac{1}{y+z},\ -\frac{x}{(y+z)^2},\ -\frac{x}{(y+z)^2}\right]\Big|_{(4,1,1)} \cdot \frac{1}{\sqrt{14}}[1,\ 2,\ 3]^T = \left[\frac{1}{2},\ -1,\ -1\right] \cdot \frac{1}{\sqrt{14}}[1,\ 2,\ 3]^T = -\frac{9}{2\sqrt{14}}$.
- The directional derivative is negative, indicating that the function decreases along the direction of $\mathbf{v}$. Based on an earlier result, we know that the maximum rate of change of a function at a point $\mathbf{x}$ is given by $||\nabla f(\mathbf{x})||$ and it is in the direction $\frac{\nabla f(\mathbf{x})}{||\nabla f(\mathbf{x})||}$.
- In the example under consideration, this maximum rate of change at $\mathbf{x}^0$ is $\frac{3}{2}$ and it is in the direction of the vector $\frac{2}{3}\left[\frac{1}{2},\ -1,\ -1\right]$.

# HW: Level Surface based Interpretation of Gradient: Examples

Let us find the maximum rate of change of the function $f(x, y, z) = x^2y^3z^4$ at the point $\mathbf{x}^0 = (1, 1, 1)$ and the direction in which it occurs. The gradient at $\mathbf{x}^0$ is $\nabla^T f\Big|_{(1,1,1)} = [2,\ 3,\ 4]$. The maximum rate of change at $\mathbf{x}^0$ is therefore $\sqrt{29}$ and the direction of the corresponding rate of change is $\frac{1}{\sqrt{29}}[2,\ 3,\ 4]$. The minimum rate of change is $-\sqrt{29}$ and the corresponding direction is $-\frac{1}{\sqrt{29}}[2,\ 3,\ 4]$.

# HW: Level Surface based Interpretation of Gradient: Examples

Determine the equations of
(a) the tangent plane to the paraboloid $\mathcal{P} : x_1 = x_2^2 + x_3^2 + 2$ at $(-1, 1, 0)$ and
(b) the normal line to the tangent plane.
To realize this as the level surface of a function of three variables, we define the function $f(x_1, x_2, x_3) = x_1 - x_2^2 - x_3^2$ and find that the paraboloid $\mathcal{P}$ is the same as the level surface $f(x_1, x_2, x_3) = -2$. The normal to the tangent plane to $\mathcal{P}$ at $\mathbf{x}^0$ is in the direction of the gradient vector $\nabla f(\mathbf{x}^0) = [1, -2, 0]^T$ and its parametric equation is
$[x_1,\ x_2,\ x_3] = [-1 + t,\ 1 - 2t,\ 0]$.
The equation of the tangent plane is therefore $(x_1 + 1) - 2(x_2 - 1) = 0$.

# Gradient and Convex Functions?

- How do we understand the behaviour of gradients for convex functions?
- While we have a lot to see in the coming sessions, here is a small peek through *sub-level sets* of a convex function

**Definition**

**[Sublevel Sets]:** Let $\mathcal{D} \subseteq \Re^n$ be a nonempty set and $f : \mathcal{D} \rightarrow \Re$. The set

$$L_\alpha(f) = \{\mathbf{x} | \mathbf{x} \in \mathcal{D},\ f(\mathbf{x}) \leq \alpha\}$$

is called the $\alpha-$sub-level set of $f$.

Now if a function $f$ is convex, the sublevel set will be convex for every value of alpha

# Gradient and Convex Functions?

- How do we understand the behaviour of gradients for convex functions?
- While we have a lot to see in the coming sessions, here is a small peek through *sub-level sets* of a convex function

**Definition**

**[Sublevel Sets]:** Let $\mathcal{D} \subseteq \Re^n$ be a nonempty set and $f : \mathcal{D} \to \Re$. The set

$$L_\alpha(f) = \left\{\mathbf{x} | \mathbf{x} \in \mathcal{D},\ f(\mathbf{x}) \leq \alpha\right\}$$

is called the $\alpha-$sub-level set of $f$.

Now if a function $f$ is convex, its $\alpha-$sub-level set is a convex set.

# Convex Function $\Rightarrow$ Convex Sub-level sets

**Theorem**

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \to \Re$ be a convex function. Then $L_\alpha(f)$ is a convex set for any $\alpha \in \Re$.*

*Proof:* Consider $\mathbf{x}_1, \mathbf{x}_2 \in L_\alpha(f)$. Then by definition of the level set, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, $f(\mathbf{x}_1) \leq \alpha$ and $f(\mathbf{x}_2) \leq \alpha$. From convexity of $\mathcal{D}$ it follows that for all $\theta \in (0,1)$, $\mathbf{x} = \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{D}$. Moreover, since $f$ is also convex,

$$f(\mathbf{x}) \leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2) \leq \theta\alpha + (1-\theta)\alpha = \alpha$$

which implies that $\mathbf{x} \in L_\alpha(f)$. Thus, $L_\alpha(f)$ is a convex set. □

Verify that for positive alpha level sets will not be conve

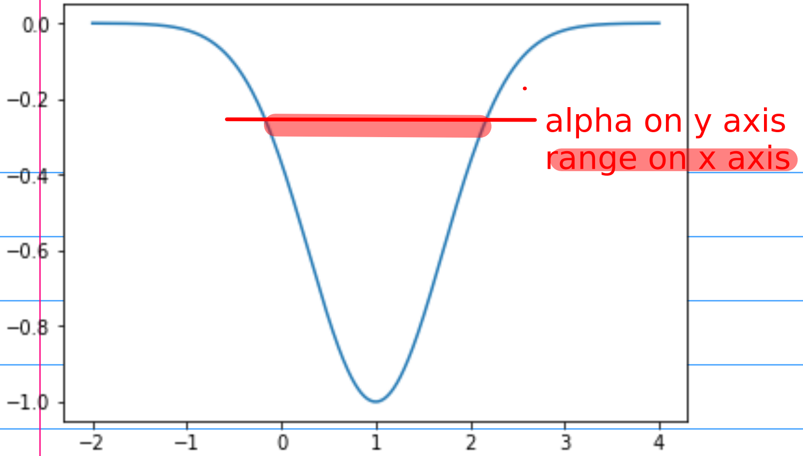
The converse of this theorem does not hold (for fixed $\alpha$ or even for all $\alpha$):

- Consider $f(\mathbf{x}) = \frac{x_2}{1+2x_1^2}$. The 0-sublevel set of this function is $\left\{(x_1, x_2) \mid x_2 \leq 0\right\}$, which is convex. However, the function $f(\mathbf{x})$ itself is not convex.

A function may be non-convex. Yet one of its sublevel sets may be con

What if all its sublevel sets were convex? Will the function be convex?

What is function is also bounded?



0.0
−0.2
−0.4
−0.6
−0.8
−1.0
−2
−1
0
1
2
3
4
alpha on y axis
range on x axis

# Convex Function $\Rightarrow$ Convex Sub-level sets

## Theorem

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \to \Re$ be a convex function. Then $L_\alpha(f)$ is a convex set for any $\alpha \in \Re$.*

*Proof:* Consider $\mathbf{x}_1, \mathbf{x}_2 \in L_\alpha(f)$. Then by definition of the level set, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$, $f(\mathbf{x}_1) \leq \alpha$ and $f(\mathbf{x}_2) \leq \alpha$. From convexity of $\mathcal{D}$ it follows that for all $\theta \in (0,1)$, $\mathbf{x} = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{D}$. Moreover, since $f$ is also convex,

$$f(\mathbf{x}) \leq \theta f(\mathbf{x}_1) + (1-\theta)f(\mathbf{x}_2) \leq \theta\alpha + (1-\theta)\alpha = \alpha$$

which implies that $\mathbf{x} \in L_\alpha(f)$. Thus, $L_\alpha(f)$ is a convex set. $\square$

The converse of this theorem does not hold (for fixed $\alpha$ or even for all $\alpha$):

- Consider $f(\mathbf{x}) = \frac{x_2}{1+2x_1^2}$. The $0$-sublevel set of this function is $\left\{(x_1, x_2) \mid x_2 \leq 0\right\}$, which is convex. However, the function $f(\mathbf{x})$ itself is not convex.
- **A function is called quasi-convex if all its sub-level sets are convex sets** Eg: Negative of the normal distribution $-\frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ is quasi-convex but not convex.

# Convex Sub-level sets $\not\Rightarrow$ Convex Function

- **A function is called quasi-convex if all its sub-level sets are convex sets**. Every quasi-convex function is not convex!
- Consider the Negative of the normal distribution $-\frac{1}{\sigma\sqrt{2\pi}}exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. This function is quasi-convex but not convex. Consider, instead, the simpler function $f(x) = -exp(-(x-\mu)^2)$.
  - Then $f'(x) =$

# Convex Sub-level sets $\not\Longrightarrow$ Convex Function

- **A function is called quasi-convex if all its sub-level sets are convex sets**. Every quasi-convex function is not convex!
- Consider the Negative of the normal distribution $-\frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. This function is quasi-convex but not convex. Consider, instead, the simpler function $f(x) = -exp(-(x-\mu)^2)$.
  - Then $f'(x) = 2(x-\mu)exp(-(x-\mu)^2)$
  - And $f''(x) = 2exp(-(x-\mu)^2) - 4(x-\mu)^2 exp(-(x-\mu)^2) = (2-4(x-\mu)^2)exp(-(x-\mu)^2)$ which is $< 0$ if $(x-\mu)^2 > \frac{1}{2}$,
  - Thus, the second derivative is negative if $x > \mu + \frac{1}{\sqrt{2}}$ or $x < -\mu - \frac{1}{\sqrt{2}}$.
  - Recall from discussion of convexity of $f : \Re \to \Re$ that if the derivative is not non-decreasing everywhere $\Longrightarrow$

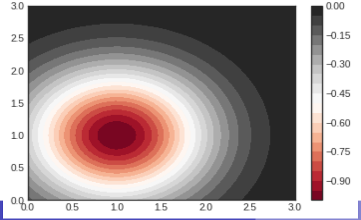
# Convex Sub-level sets $\not\Longrightarrow$ Convex Function

- **A function is called quasi-convex if all its sub-level sets are convex sets**. Every quasi-convex function is not convex!
- Consider the Negative of the normal distribution $-\frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$. This function is quasi-convex but not convex. Consider, instead, the simpler function $f(x) = -exp(-(x-\mu)^2)$.
  - Then $f'(x) = 2(x-\mu)exp(-(x-\mu)^2)$
  - And $f''(x) = 2exp(-(x-\mu)^2) - 4(x-\mu)^2 exp(-(x-\mu)^2) = (2-4(x-\mu)^2)exp(-(x-\mu)^2)$ which is $< 0$ if $(x-\mu)^2 > \frac{1}{2}$,
  - Thus, the second derivative is negative if $x > \mu + \frac{1}{\sqrt{2}}$ or $x < -\mu - \frac{1}{\sqrt{2}}$.
  - Recall from discussion of convexity of $f : \Re \rightarrow \Re$ that if the derivative is not non-decreasing everywhere $\Longrightarrow$ function is not convex everywhere.
- To prove that this function is quasi-convex, we can ....

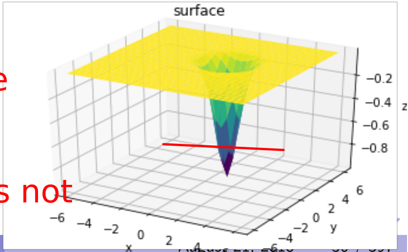
# Proof that the function is Quasi-Convex

1. Inspect the $L_\alpha(f)$ sublevel sets of this function:
   $L_\alpha(f) = \{x| - exp(-(x-\mu)^2) \le \alpha\} = \{x|exp(-(x-\mu)^2) \ge -\alpha\}$.
2. Since $exp(-(x-\mu)^2)$ is monotonically increasing for $x < \mu$ and monotonically decreasing for $x > \mu$, the set $\{x|exp(-(x-\mu)^2) \ge -\alpha\}$ will be a contiguous closed interval around $\mu$ and therefore a convex set.
3. Thus, $f(x) = -exp(-(x-\mu)^2)$ is quasi-convex (and so is its generalization - the negative of the normal density function).

One can similarly prove that the negative of the multivariate normal density function $f(\mathbf{x}) = -\frac{1}{\sqrt{|\Sigma|(2\pi)^n}} exp\left(-(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)\right)$ is also quasi-convex:




The sublevel sets in R2 are all ellipsoids
The function graph in R3 is not convex



# Proof that the function is Quasi-Convex

1. Inspect the $L_\alpha(f)$ sublevel sets of this function:
   $L_\alpha(f) = \{x | -exp(-(x-\mu)^2) \leq \alpha\} = \{x | exp(-(x-\mu)^2) \geq -\alpha\}$.
2. Since $exp(-(x-\mu)^2)$ is monotonically increasing for $x < \mu$ and monotonically decreasing for $x > \mu$, the set $\{x | exp(-(x-\mu)^2) \geq -\alpha\}$ will be a contiguous closed interval around $\mu$ and therefore a convex set.
3. Thus, $f(x) = -exp(-(x-\mu)^2)$ is quasi-convex (and so is its generalization - the negative of the normal density function).

One can similarly prove that the negative of the multivariate normal density function $f(\mathbf{x}) = -\frac{1}{\sqrt{|\Sigma|(2\pi)^n}} exp\left(-(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right)$ is also quasi-convex:

$$L_\alpha(f) = \left\{\mathbf{x} \middle| -exp\left(-(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right) \leq \alpha\sqrt{|\Sigma|(2\pi)^n}\right\} =$$

$$\left\{\mathbf{x} \middle| exp\left(-(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right) \geq -\alpha\sqrt{|\Sigma|(2\pi)^n}\right\} =$$

$$\left\{\mathbf{x} \middle| \left((\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)\right) \leq -\log\left(-\alpha\sqrt{|\Sigma|(2\pi)^n}\right)\right\}$$ which is an **ellipsoid**. Verify!

# Quasi-Convex Functions and Optimization

- Consider a minimization problem with a quasi-convex objective $q(\mathbf{x})$ and convex functions $f_1(\mathbf{x})...f_m(\mathbf{x})$ in the constraints

Eg: maximizing likelihood of gaussian fits is equivalent to this

$$\begin{aligned} \text{minimize} \quad & q(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) \leq 0 \quad \text{for each } i = 1..m \end{aligned} \tag{4}$$

We note that the constraint set is intersection over the 0 sublevel sets of the fi's.

# Quasi-Convex Functions and Optimization

- Consider a minimization problem with a quasi-convex objective $q(\mathbf{x})$ and convex functions $f_1(\mathbf{x})...f_m(\mathbf{x})$ in the constraints

$$\begin{array}{lll} \text{minimize} & q(\mathbf{x}) & \\ \text{subject to} & f_i(\mathbf{x}) \leq 0 & \text{for each } i = 1..m \end{array} \tag{4}$$

We note that the constraint set is convex since (i) each $f_i(\mathbf{x}) \leq$ is convex sub-level set of a convex function $f_i(\mathbf{x})$ and (ii) intersection of finite convex sets is convex.

# Quasi-Convex Functions and Optimization

- Consider a minimization problem with a quasi-convex objective $q(\mathbf{x})$ and convex functions $f_1(\mathbf{x})...f_m(\mathbf{x})$ in the constraints

$$\begin{aligned} &\text{minimize} && q(\mathbf{x}) \\ &\text{subject to} && f_i(\mathbf{x}) \le 0 \quad \text{for each } i = 1..m \end{aligned} \tag{4}$$

We note that the constraint set is convex since (i) each $f_i(\mathbf{x}) \le$ is convex sub-level set of a convex function $f_i(\mathbf{x})$ and (ii) intersection of finite convex sets is convex.

- How do we proceed from a quasi-convex $q(\mathbf{x})$ to complete convexity? Consider:

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && q(\mathbf{x}) \le t \\ &\text{and} && f_i(\mathbf{x}) \le 0 \quad \text{for each } i = 1..m \end{aligned} \tag{5}$$

linear function in objective is convex

convex constraint set

This is a problem with convex objective and convex constraint set

# Quasi-Convex Functions and Optimization

- Consider a minimization problem with a quasi-convex objective $q(\mathbf{x})$ and convex functions $f_1(\mathbf{x})...f_m(\mathbf{x})$ in the constraints

$$\begin{aligned} &\text{minimize} && q(\mathbf{x}) \\ &\text{subject to} && f_i(\mathbf{x}) \leq 0 \quad \text{for each } i = 1..m \end{aligned} \tag{4}$$

  We note that the constraint set is convex since (i) each $f_i(\mathbf{x}) \leq$ is convex sub-level set of a convex function $f_i(\mathbf{x})$ and (ii) intersection of finite convex sets is convex.
- How do we proceed from a quasi-convex $q(\mathbf{x})$ to complete convexity? Consider:

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && q(\mathbf{x}) \leq t \\ &\text{and} && f_i(\mathbf{x}) \leq 0 \quad \text{for each } i = 1..m \end{aligned} \tag{5}$$

  can be posed as

  This is a convex feasibility problem (convex objective and convex constraint set) and can be solved as a series of

  bisection search on convex feasibility

# Quasi-Convex Functions and Optimization

- Consider a minimization problem with a quasi-convex objective $q(\mathbf{x})$ and convex functions $f_1(\mathbf{x})...f_m(\mathbf{x})$ in the constraints

$$\begin{aligned} &\text{minimize} && q(\mathbf{x}) \\ &\text{subject to} && f_i(\mathbf{x}) \leq 0 \quad \text{for each } i = 1..m \end{aligned} \tag{4}$$

  We note that the constraint set is convex since (i) each $f_i(\mathbf{x}) \leq$ is convex sub-level set of a convex function $f_i(\mathbf{x})$ and (ii) intersection of finite convex sets is convex.

- How do we proceed from a quasi-convex $q(\mathbf{x})$ to complete convexity? Consider:

Not the most brilliant way to optimize for gaussian likelihood!

$$\begin{aligned} &\text{minimize} && t \\ &\text{subject to} && q(\mathbf{x}) \leq t \\ &\text{and} && f_i(\mathbf{x}) \leq 0 \quad \text{for each } i = 1..m \end{aligned} \tag{5}$$

what if we take log of gaussian?

is it concave (its negative convex)

  This is a convex feasibility problem (convex objective and convex constraint set) and can be solved as a series of convex (feasibility) optimization problems using bisection search on $t$ (see Section 4.2.5 of Boyd and Vandenberghe)

In general refer to 4.2.5 of Boyd for operations that preserve quasi-convexity
And what about operations that convert quasi-convex function into a convex function? -Log(-f(x)) ?

# Gradient, Convex Functions and Sub-level sets: A First Peek

We have already seen that

- The gradient $\nabla f(\mathbf{x}^*)$ at $\mathbf{x}^*$ is normal to the tangent hyperplane to the level set $\{\mathbf{x} | f(\mathbf{x}) = f(\mathbf{x}^*)\}$ at $\mathbf{x}^*$
- The gradient $\nabla f(\mathbf{x}^*)$ at $\mathbf{x}^*$ points in direction of increasing values of $f(.)$ at $\mathbf{x}^*$

independent of convexity of f

Now, if $f(\mathbf{x})$ is also convex

the gradient gives you a tangential hyperplane that is a supporting hyperplane to the sublevel set at that point

# Gradient, Convex Functions and Sub-level sets: A First Peek

We have already seen that

- The gradient $\nabla f(\mathbf{x}^*)$ at $\mathbf{x}^*$ is normal to the tangent hyperplane to the level set $\{\mathbf{x}|f(\mathbf{x}) = f(\mathbf{x}^*)\}$ at $\mathbf{x}^*$
- The gradient $\nabla f(\mathbf{x}^*)$ at $\mathbf{x}^*$ points in direction of increasing values of $f(.)$ at $\mathbf{x}^*$

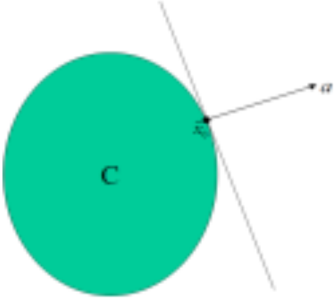
Now, if $f(\mathbf{x})$ is also convex

- The gradient $\nabla f(\mathbf{x}^*)$ **at $\mathbf{x}^*$ is normal to the tangent hyperplane to the sub-level set $L_{f(\mathbf{x}^*)}(f) = \{\mathbf{x}|f(\mathbf{x}) \leq f(\mathbf{x}^*)\}$ at $\mathbf{x}^*$**, pointing away from the set $L_{f(\mathbf{x}^*)}(f)$
- The tangent hyperplane defined by $\nabla f(\mathbf{x}^*)$ at $\mathbf{x}^*$ is a **supporting hyperplane** to the convex set $\{\mathbf{x}|f(\mathbf{x}) \leq f(\mathbf{x}^*)\}$ at $\mathbf{x}^*$

# Recall: Supporting hyperplane and Convex Sets

**Supporting hyperplane** to set $\mathcal{C}$ at boundary point $\mathbf{x}_o$:

- $\left\{\mathbf{x}|\mathbf{a}^T\mathbf{x} = \mathbf{a}^T\mathbf{x}_o\right\}$
- where $\mathbf{a} \neq 0$ and $\mathbf{a}^T\mathbf{x} \leq \mathbf{a}^T\mathbf{x}_o$ for all $\mathbf{x} \in \mathcal{C}$




H/W: Are sublevel sets always closed? Do they contain the boundary point?

**Recall Supporting hyperplane theorem:** if $\mathcal{C}$ is convex, then there exists a supporting hyperplane at every boundary point of C.

# Convex Functions and Their Epigraphs

We saw that a convex function has a convex sub-level set. But the converse is not true. Is there a set corresponding to a function such that one is convex if and only if the other is?

YES: Set of points lying above the graph of the function
Also called "Epigraph"

# Convex Functions and Their Epigraphs

We saw that a convex function has a convex sub-level set. But the converse is not true. Is there a set corresponding to a function such that one is convex if and only if the other is?

## Definition

**[Epigraph]:** Let $\mathcal{D} \subseteq \Re^n$ be a nonempty set and $f : \mathcal{D} \to \Re$. The set $\{(\mathbf{x}, f(\mathbf{x}) | \mathbf{x} \in \mathcal{D}\}$ is called graph of $f$ and lies in $\Re^{n+1}$. The epigraph of $f$ is a subset of $\Re^{n+1}$ and is defined as

$$epi(f) = \{(\mathbf{x}, \alpha) | f(\mathbf{x}) \leq \alpha, \ \mathbf{x} \in \mathcal{D}, \ \alpha \in \Re\} \tag{6}$$

**In some sense, the epigraph is the set of points lying above the graph of $f$.**

Eg: Recall affine functions of vectors: $\mathbf{a}^T\mathbf{x} + b$ where $\mathbf{a} \in \Re^n$. Its epigraph is $\{(\mathbf{x}, t) | \mathbf{a}^T\mathbf{x} + b \leq t\} \subseteq \Re^{n+1}$ which is a half-space (a convex set).

# Convex Functions and Their Epigraphs

## Definition

**[Hypograph]:** Similarly, the *hypograph* of $f$ is a subset of $\Re^{n+1}$, lying below the graph of $f$ and is defined by

$$hyp(f) = \left\{(\mathbf{x}, \alpha) | f(\mathbf{x}) \geq \alpha,\ \mathbf{x} \in \mathcal{D},\ \alpha \in \Re\right\} \tag{7}$$

f is concave function if and only if its hypograph is convex set

# Convex Functions and Their Epigraphs (contd)

There is a one to one correspondence between the convexity of function $f$ and that of the set $epi(f)$, as stated in the following result.

**Theorem**

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \to \Re$. Then*

# Convex Functions and Their Epigraphs (contd)

There is a one to one correspondence between the convexity of function $f$ and that of the set $epi(f)$, as stated in the following result.

**Theorem**

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \rightarrow \Re$. Then $f$ is convex if and only if $epi(f)$ is a convex set.*

*Proof:* $f$ **convex function** $\implies$ $epi(f)$ **convex set**

Proof has similar traits as proof for sublevel sets

# Convex Functions and Their Epigraphs (contd)

There is a one to one correspondence between the convexity of function $f$ and that of the set $epi(f)$, as stated in the following result.

**Theorem**

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \to \Re$. Then $f$ is convex if and only if $epi(f)$ is a convex set.*

*Proof:* $f$ **convex function** $\implies$ $epi(f)$ **convex set**

Let $f$ be convex. For any $(\mathbf{x}_1, \alpha_1) \in epi(f)$ and $(\mathbf{x}_2, \alpha_2) \in epi(f)$ and any $\theta \in (0, 1)$,

$$f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2)) \leq$$ use property of membership above

By convexity of f

# Convex Functions and Their Epigraphs (contd)

There is a one to one correspondence between the convexity of function $f$ and that of the set $epi(f)$, as stated in the following result.

**Theorem**

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \to \Re$. Then $f$ is convex if and only if $epi(f)$ is a convex set.*

*Proof:* $f$ **convex function** $\implies$ $epi(f)$ **convex set**

Let $f$ be convex. For any $(\mathbf{x}_1, \alpha_1) \in epi(f)$ and $(\mathbf{x}_2, \alpha_2) \in epi(f)$ and any $\theta \in (0, 1)$,

$$f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1-\theta)f(\mathbf{x}_2)) \leq \theta\alpha_1 + (1-\theta)\alpha_2$$

Since $\mathcal{D}$ is convex, $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{D}$. Therefore,
$(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta\alpha_1 + (1-\theta)\alpha_2) \in$ epi(f)

# Convex Functions and Their Epigraphs (contd)

There is a one to one correspondence between the convexity of function $f$ and that of the set $epi(f)$, as stated in the following result.

**Theorem**

*Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \to \Re$. Then $f$ is convex if and only if $epi(f)$ is a convex set.*

*Proof:* $f$ **convex function** $\implies$ $epi(f)$ **convex set**

Let $f$ be convex. For any $(\mathbf{x}_1, \alpha_1) \in epi(f)$ and $(\mathbf{x}_2, \alpha_2) \in epi(f)$ and any $\theta \in (0, 1)$,

$$f(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2)) \leq \theta \alpha_1 + (1-\theta)\alpha_2$$

Since $\mathcal{D}$ is convex, $\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{D}$. Therefore, $\left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta\alpha_1 + (1-\theta)\alpha_2\right) \in epi(f)$. Thus, $epi(f)$ is convex if $f$ is convex. This proves the necessity part.

# Convex Functions and Their Epigraphs (contd)

***epi*($f$) convex set $\implies$ $f$ convex function**

To prove sufficiency, assume that *epi*($f$) is convex. Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$. So, $\left(\mathbf{x}_1, f(\mathbf{x}_1)\right) \in epi(f)$ and $\left(\mathbf{x}_2, f(\mathbf{x}_2)\right) \in epi(f)$. Since *epi*($f$) is convex, for $\theta \in (0,1)$,

$$\left(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta\alpha_1 + (1-\theta)\alpha_2\right) \in epi(f)$$

which implies that f must also be convex!

# Convex Functions and Their Epigraphs (contd)

*epi*$(f)$ **convex set** $\implies$ $f$ **convex function**

To prove sufficiency, assume that *epi*$(f)$ is convex. Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}$. So, $\big(\mathbf{x}_1, f(\mathbf{x}_1)\big) \in$ *epi*$(f)$ and $\big(\mathbf{x}_2, f(\mathbf{x}_2)\big) \in$ *epi*$(f)$. Since *epi*$(f)$ is convex, for $\theta \in (0,1)$,

$$\big(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta\alpha_1 + (1-\theta)\alpha_2\big) \in epi(f)$$

which implies that $f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1-\theta)f(\mathbf{x}_2))$ for any $\theta \in (0,1)$. This proves the sufficiency. $\square$

## Epigraph and Convexity

- Given a convex function $f(\mathbf{x})$ and a convex domain $\mathcal{D}$, the convex optimization problem

$$\min_{\mathbf{x}\in\mathcal{D}} f(\mathbf{x})$$

can be equivalently expressed as

$$\min_{\mathbf{x}\in\mathcal{D}, t\in\Re, f(\mathbf{x})\leq t} t =$$

# Epigraph and Convexity

- Given a convex function $f(\mathbf{x})$ and a convex domain $\mathcal{D}$, the convex optimization problem

$$\min_{\mathbf{x}\in\mathcal{D}} f(\mathbf{x})$$

can be equivalently expressed as

$$\min_{\mathbf{x}\in\mathcal{D},t\in\Re,\underline{f(\mathbf{x})\leq t}} t = \min_{\mathbf{x}\in\mathcal{D},\underline{(\mathbf{x},t)\in epi(f)}} t$$

minimize upper bound on f

- Recall the first order condition for convexity of a differentiable function $f:\Re\to\Re$. Is there an equivalent for $f:\mathcal{D}\to\Re$?

Key idea: Supporting hyperplane to epigraph is
The lower bound to the graph

# Epigraph and Convexity

- Given a convex function $f(\mathbf{x})$ and a convex domain $\mathcal{D}$, the convex optimization problem

$$\min_{\mathbf{x}\in\mathcal{D}} f(\mathbf{x})$$

can be equivalently expressed as

$$\min_{\mathbf{x}\in\mathcal{D},t\in\Re,f(\mathbf{x})\leq t} t = \min_{\mathbf{x}\in\mathcal{D},(\mathbf{x},t)\in epi(f)} t$$

- Recall the first order condition for convexity of a differentiable function $f:\Re\to\Re$. Is there an equivalent for $f:\mathcal{D}\to\Re$? Let $f:\mathcal{D}\to\Re$ be a differentiable convex function on an open convex set $\mathcal{D}$. Then $f$ is convex if and only if, for any $\mathbf{x},\mathbf{y}\in\mathcal{D}$,

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

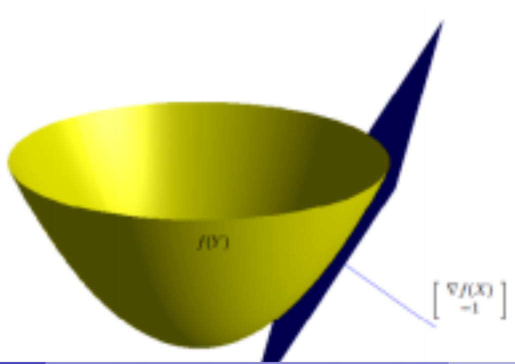
First order taylor expansion lower bounds

# Epigraph, Convexity and Gradients

..(contd).... $f$ is convex if and only if, for any $\mathbf{x}, \mathbf{y} \in \mathcal{D}$,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla^T f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \tag{8}$$

If $\mathcal{D} \subseteq \Re^n$, this means that for each and every point $\mathbf{x} \in \mathcal{D}$ for a convex real function $f(\mathbf{x})$, there exists a hyperplane $H \in \Re^{n+1}$ having normal $[\nabla f(\mathbf{x}) \ \ -1]^T$ supporting the function epigraph at $[\mathbf{x} \ \ f(\mathbf{x})]^T$. See Figure below sourced from https://ccrma.stanford.edu/~dattorro/gcf.pdf



# Epigraph, Convexity, Gradients and Level-sets

- **Revisiting level sets:** We can embed the graph of a function of $n$ variables as the $0$-level set of a function of $n+1$ variables

[1](that is, the tangent hyperplane to $f(\mathbf{x})$ at the point $\mathbf{x}$)

# Epigraph, Convexity, Gradients and Level-sets

- **Revisiting level sets:** We can embed the graph of a function of $n$ variables as the $0$-level set of a function of $n+1$ variables
- More concretely, if $f : \mathcal{D} \rightarrow \Re,\ \mathcal{D} \subseteq \Re^n$ then we define $F : \mathcal{D}' \rightarrow \Re,\ \mathcal{D}' = \mathcal{D} \times \Re$ as $F(\mathbf{x}, z) = f(\mathbf{x}) - z$ with $\mathbf{x} \in \mathcal{D}'$.
- The gradient of $F$ at any point $(\mathbf{x}, z)$ is simply, $\nabla F(\mathbf{x}, z) = \left[f_{x_1}, f_{x_2}, \ldots, f_{x_n}, -1\right]$ with the first $n$ components of $\nabla F(\mathbf{x}, z)$ given by the $n$ components of $\nabla f(\mathbf{x})$.
- The graph of $f$ can be recovered as the $0-$level set of $F$ given by $F(\mathbf{x}, z) = 0$.
- The equation of the tangent hyperplane $(\mathbf{y}, z)$ to the $0-$level set of $F$ at the point $(\mathbf{x}, f(\mathbf{x}))$ is[1] $\nabla^T F(\mathbf{x}, f(\mathbf{x})).[\mathbf{y} - \mathbf{x}, z - f(\mathbf{x})]^T = [\nabla f(\mathbf{x}), -1]^T.[\mathbf{y} - \mathbf{x}, z - f(\mathbf{x})]^T = 0$.

[1](that is, the tangent hyperplane to $f(\mathbf{x})$ at the point $\mathbf{x}$)

# Epigraph, Convexity, Gradients and Level-sets (contd.)

Substituting appropriate expression for $\nabla F(\mathbf{x})$, the equation of the tangent plane $(\mathbf{y}, z)$ can be written as

$$\left(\sum_{i=1}^{n} f_{x_i}(\mathbf{x})(y_i - x_i)\right) - \big(z - f(\mathbf{x})\big) = 0$$

or equivalently as,

$$\left(\nabla^T f(\mathbf{x})(\mathbf{y} - \mathbf{x})\right) + f(\mathbf{x}) = z$$

# Epigraph, Convexity, Gradients and Level-sets (contd.)

Substituting appropriate expression for $\nabla F(\mathbf{x})$, the equation of the tangent plane $(\mathbf{y}, z)$ can be written as

$$\left(\sum_{i=1}^{n} f_{x_i}(\mathbf{x})(y_i - x_i)\right) - \big(z - f(\mathbf{x})\big) = 0$$

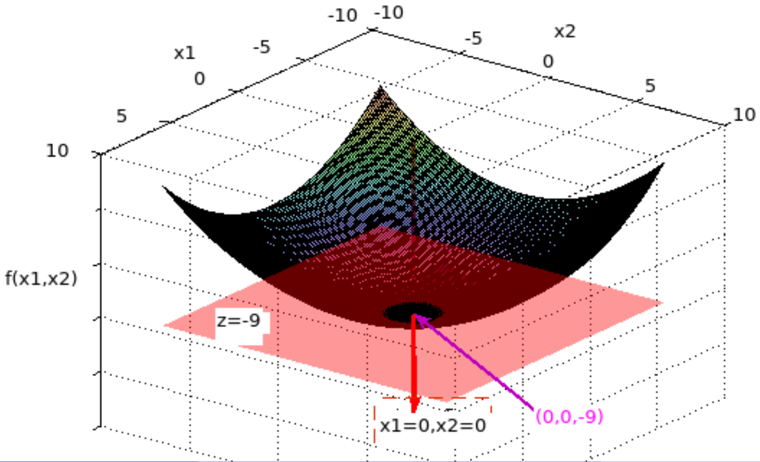
or equivalently as,

$$\left(\nabla^T f(\mathbf{x})(\mathbf{y} - \mathbf{x})\right) + f(\mathbf{x}) = z$$

Revisiting the gradient-based condition for convexity in (8), we have that for a convex function, $f(\mathbf{y})$ is greater than each such $z$ on the hyperplane: $f(\mathbf{y}) \geq z = f(\mathbf{x}) + \nabla^T f(\mathbf{x})(\mathbf{y} - \mathbf{x})$

# Gradient and Epigraph (contd)

As an example, consider the paraboloid, $f(x_1, x_2) = x_1^2 + x_2^2 - 9$ that attains its minimum at $(0, 0)$. We see below its epigraph.



# Illustrations to understand Gradient

- For the paraboloid, $f(x_1, x_2) = x_1^2 + x_2^2 - 9$, the corresponding $F(x_1, x_2, z) = x_1^2 + x_2^2 - 9 - z$ and the point $x^0 = (\mathbf{x}^0, z) = (1, 1, -7)$ which lies on the 0-level surface of $F$. The gradient $\nabla F(x_1, x_2, z)$ is $[2x_1,\ 2x_2,\ -1]$, which when evaluated at $x^0 = (1, 1, -7)$ is $[-2,\ -2,\ -1]$. The equation of the tangent plane to $f$ at $x^0$ is therefore given by $2(x_1 - 1) + 2(x_2 - 1) - 7 = z$.
- The paraboloid attains its minimum at $(0, 0)$. Plot the tanget plane to the surface at $(0, 0, f(0, 0))$ as also the gradient vector $\nabla F$ at $(0, 0, f(0, 0))$. What do you expect?