Convexity, Local and Global Optimality, etc.

Optimization Principles for Multivariate Functions

In the following, we state some important properties of convex functions, some of which require knowledge of 'derivatives' in $\Re^n$. These also include relationships between convex functions and convex sets, and first and second order conditions for convexity.

Convex Function $\Rightarrow$ Convex Sub-level sets

Theorem

Let $\mathcal{D} \subseteq \Re^n$ be a nonempty convex set, and $f : \mathcal{D} \rightarrow \Re$ 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. To illustrate this, consider the function $f(\mathbf{x}) = \frac{x_2}{1+2x_1^2}$. 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. [Homework]

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 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 $f : \Re \rightarrow \Re$ if the derivative is not non-decreasing everywhere $\implies$ 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:
   $L_\alpha(f) = \{x| - exp(-(x-\mu)^2) \leq \alpha\} = \{x|exp(-(x-\mu)^2) \geq -\alpha\}$.
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 μ 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 is also quasi-convex, by inspecting its sub-level sets, which are nothing but **ellipsoids**.

Every convex function is a Quasi-convex function
Not every quasi-convex function is a convex function

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 μ 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 is also quasi-convex, by inspecting its sub-level sets, which are nothing but **ellipsoids**.

Note that

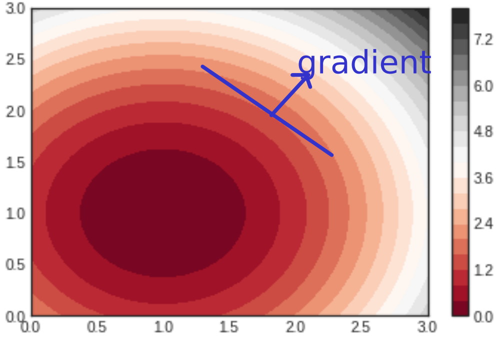
Gradient, Convex Functions and Sub-level sets: A First Peek

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

At every point x* the gradient at x* gives you a hyperplane that supports the sublevel set {x | f(x) <= f(x*)}



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

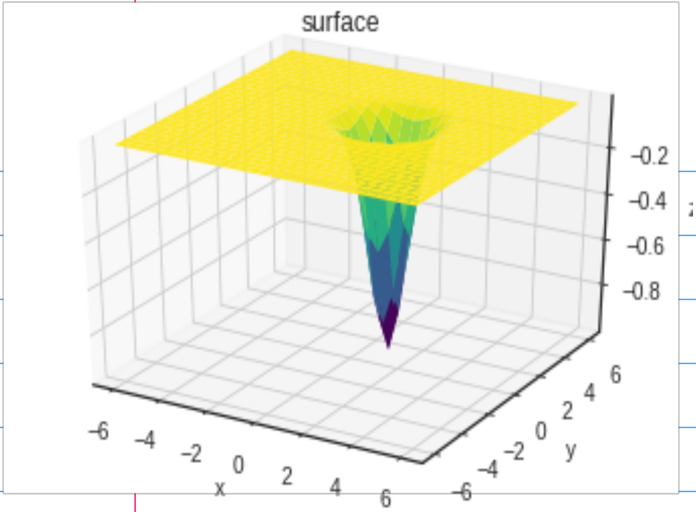
We have already seen that

- 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 $\mathbf{x}^*$

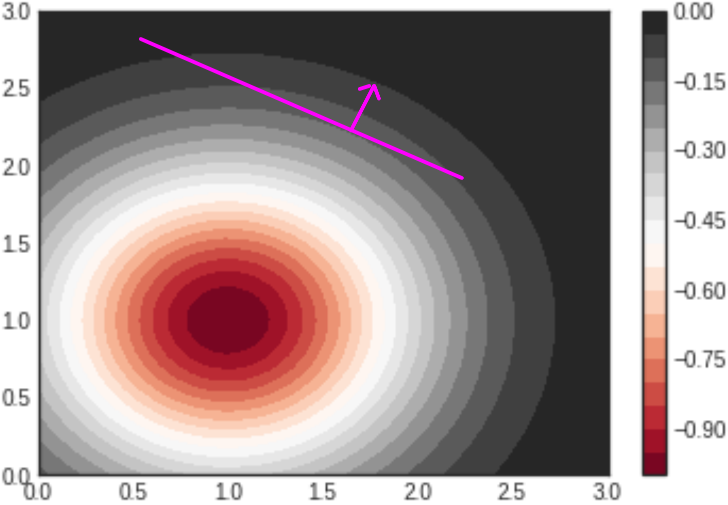
Now, if $f(\mathbf{x})$ is also convex

- 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 $\mathbf{x}^*$

As we have seen, this also holds for quasi convex functions such as negative of normal density function (see next page)



Non convex function
Negative of normal density



But it is quasi-convex owing to convex subleve sets

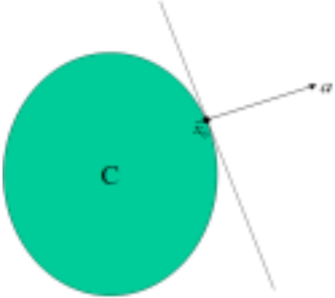
Hence the gradient continues to give us a supporting hyperpla

Aside: 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 $\mathbf{x} \in \mathcal{C}$

Supporting because entire convex set lies on one side of half space



If the set C is not closed, the point xo will not lie in C

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

Let us further the connection between convex functions and sets by introducing the concept of the *epigraph* of a function.

Definition

[Epigraph]: Let $\mathcal{D} \subseteq \Re^n$ be a nonempty set and $f : \mathcal{D} \to \Re$. The set $\left\{(\mathbf{x}, f(\mathbf{x})|\mathbf{x} \in \mathcal{D}\right\}$ 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

$$epi(f) = \left\{(\mathbf{x}, \alpha)|f(\mathbf{x}) \leq \alpha,\ \mathbf{x} \in \mathcal{D},\ \alpha \in \Re\right\} \tag{5}$$

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).

Here a convex function has a convex epigraph. Is that always true?

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

$$hyp(f) = \left\{(\mathbf{x}, \alpha) | f(\mathbf{x}) \geq \alpha, \ \mathbf{x} \in \mathcal{D}, \ \alpha \in \Re\right\} \tag{6}$$

Hypographs are interesting for concave functions

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 convex

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**

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**

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

Is the epi(f) of a quasi-convex f also always convex

NO: Not always. It is true only if the f is convex.

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 of t wrt (x,t) on epi(f)

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 $f : \Re \rightarrow \Re$. Is there an equivalent for $f : \mathcal{D} \rightarrow \Re$?

f'(x) is (strictly) increasing OR

f at y obtained using linear approx at x is (strictly) less than or equal to the actual value f(y) at y

f(x) + Direc deriv along y-x <= f(y)

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

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

- 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 $\mathbf{x},\mathbf{y}\in\mathcal{D}$,

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

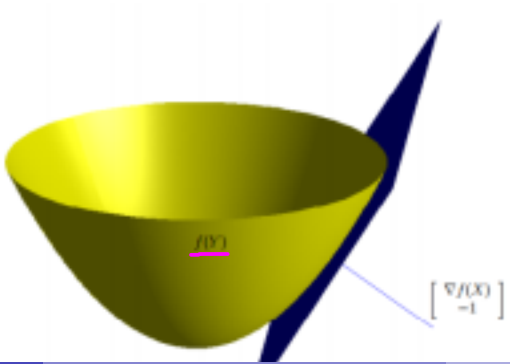
directional derivative at x along y-x

Epigraph, Convexity and Gradients

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

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

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 $[\nabla f(\mathbf{x}) \quad -1]^T$ supporting the function epigraph at $[\mathbf{x} \ \ f(\mathbf{x})]^T$. See Figure below sourced from https://ccrma.stanford.edu/~dattorro/gcf.pdf



Homework: Equation of Tangent Hyperplane to Epigraph

For the function $f: \mathcal{D} \rightarrow \Re$ such that $\mathcal{D} \subseteq \Re^n$, write down the equation of the tangent hyperplane to its epigraph at $(\mathbf{x}^0, f(\mathbf{x}^0))$.
As an example, consider the paraboloid, $f(x_1, x_2) = x_1^2 + x_2^2 - 9$ and $\mathbf{x}^0 = (0, 0)$ and write the equation of its tangent hyperplane. Also try and plot.

First-Order Convexity Conditions: The complete statement

Theorem

1. *For differentiable $f : \mathcal{D} \rightarrow \Re$ and open convex set $\mathcal{D}$, f is convex* **iff**, *for any $\mathbf{x}, \mathbf{y} \in \mathcal{D}$,*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla^T f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \tag{7}$$

2. *f is strictly convex* **iff**, *for any $\mathbf{x}, \mathbf{y} \in \mathcal{D}$, with $\mathbf{x} \neq \mathbf{y}$,*

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

3. *f is strongly convex* **iff**, *for any $\mathbf{x}, \mathbf{y} \in \mathcal{D}$, and for some constant $c > 0$,*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla^T f(\mathbf{x})(\mathbf{y} - \mathbf{x}) + \frac{1}{2} c ||\mathbf{y} - \mathbf{x}||^2 \tag{9}$$