Composition with scalar functions

composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$:

$$f(x) = h(g(x))$$

f is convex if h is convex, g is convex & h is non-decreasing

- proof $f''(x) = h''(g(x))g'(x)^2 + g''(x)h'(g(x))$ (n=1)

$$\nabla^2 f(x) = h''(g(x))\nabla g(x)\nabla^T g(x) + \nabla^2 g(x) h'(g(x))$$

(Assume differentiability)

Another option: h is convex, g is concave & h is non-increasing

examples

(H/W: Argue based on first principles / defn of convexity)

$f(x) = \sum \log(x_i)$ ✓ is concave. (See next page: $g_i(x) = x_i$)

$f(x) = \log(\sum x_i)$ (H/W)

Composition with scalar functions

composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$:

$$f(x) = h(g(x))$$

f is convex if $\begin{array}{l} g \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing} \\ g \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing} \end{array}$

- proof (for $n = 1$, differentiable g, h)

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$$

- note: monotonicity must hold for extended-value extension $\tilde{h}$

examples

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive

So that non-decreasing

Vector composition

composition of $g : \mathbf{R}^n \to \mathbf{R}^k$ and $h : \mathbf{R}^k \to \mathbf{R}$:

$$f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_k(x))$$

f is convex if $\begin{array}{l} g_i \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing in each argument} \\ g_i \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing in each argument} \end{array}$

proof (for $n = 1$, differentiable g, h)

Each component of $\nabla h \geq 0$

Hessian $= n^2 \times k$

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

$n \times n$ Jacobian = Matrix of gradients $= n \times k$

examples

- $\sum_{i=1}^m \log g_i(x)$ is concave if g_i are concave and positive
- $\log \sum_{i=1}^m \exp g_i(x)$ is convex if g_i are convex

Note: Non-negative mixture of psd matrices is psd

More examples:

Theorem 78 *Let $f : \mathcal{D} \rightarrow \Re$ with $\mathcal{D} \subseteq \Re^n$ be differentiable on the convex set $\mathcal{D}$. Then,*

1. *f is convex on $\mathcal{D}$ if and only if is its gradient ∇f is monotone. That is, for all $\mathbf{x}, \mathbf{y} \in \Re$*

$$\left(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\right)^T (\mathbf{x} - \mathbf{y}) \geq 0 \tag{4.53}$$

For 1-d: If $x \geq y$ then this means $\nabla f(x) \geq \nabla f(y)$

2. *f is strictly convex on $\mathcal{D}$ if and only if is its gradient ∇f is strictly monotone. That is, for all $\mathbf{x}, \mathbf{y} \in \Re$ with $\mathbf{x} \neq \mathbf{y}$,*

$$\left(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\right)^T (\mathbf{x} - \mathbf{y}) > 0 \tag{4.54}$$

3. *f is uniformly or strongly convex on $\mathcal{D}$ if and only if is its gradient ∇f is uniformly monotone. That is, for all $\mathbf{x}, \mathbf{y} \in \Re$,*

$$\left(\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\right)^T (\mathbf{x} - \mathbf{y}) \geq c||\mathbf{x} - \mathbf{y}||^2 \tag{4.55}$$

for some constant $c > 0$.

Necessity: Suppose f is uniformly convex on $\mathcal{D}$. Then from theorem 75, we know that for any $\mathbf{x}, \mathbf{y} \in \mathcal{D}$,

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

Adding the two inequalities, we get (4.55). If f is convex, the inequalities hold with $c = 0$, yielding (4.54). If f is strictly convex, the inequalities will be strict, yielding (4.54).

Sufficiency: Suppose ∇f is monotone. For any fixed $\mathbf{x}, \mathbf{y} \in \mathcal{D}$, consider the function $\phi(t) = f(\mathbf{x} + t(\mathbf{y}-\mathbf{x}))$. By the mean value theorem applied to $\phi(t)$, we should have for some $t \in (0,1)$,

$$\phi(1) - \phi(0) = \phi'(t) \tag{4.56}$$

Letting $\mathbf{z} = \mathbf{x} + t(\mathbf{y}-\mathbf{x})$, (4.56) translates to

$$f(\mathbf{y}) - f(\mathbf{x}) = \nabla^T f(\mathbf{z})(\mathbf{y}-\mathbf{x}) \tag{4.57}$$

Also, by definition of monotonicity of ∇f, (from (4.53)),

$$\left(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})\right)^T (\mathbf{y}-\mathbf{x}) = \frac{1}{t}\left(\nabla f(\mathbf{z}) - \nabla f(\mathbf{x})\right)^T (\mathbf{z}-\mathbf{x}) \geq 0 \tag{4.58}$$

Combining (4.57) with (4.58), we get,

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &= \left(\nabla f(\mathbf{z}) - f(\mathbf{x})\right)^T (\mathbf{y}-\mathbf{x}) + \nabla^T f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \\ &\geq \nabla^T f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \end{aligned} \tag{4.59}$$

By theorem 75, this inequality proves that f is convex. Strict convexity can be similarly proved by using the strict inequality in (4.58) inherited from strict monotonicity, and letting the strict inequality follow through to (4.59). For the case of strong convexity, from (4.55), we have

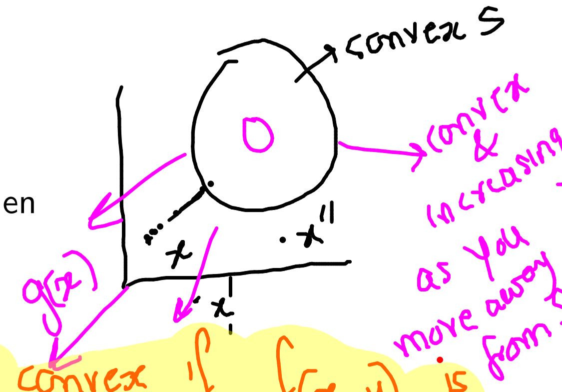
$$\begin{aligned} \phi'(t) - \phi'(0) &= \left(\nabla f(\mathbf{z}) - f(\mathbf{x})\right)^T (\mathbf{y}-\mathbf{x}) \\ = \frac{1}{t}\left(\nabla f(\mathbf{z}) - f(\mathbf{x})\right)^T (\mathbf{z}-\mathbf{x}) &\geq \frac{1}{t}c||\mathbf{z}-\mathbf{x}||^2 = ct||\mathbf{y}-\mathbf{x}||^2 \end{aligned} \tag{4.60}$$

$$\phi(1) - \phi(0) - \phi'(0) = \int_0^1 [\phi'(t) - \phi'(0)]dt \geq \frac{1}{2}c||\mathbf{y}-\mathbf{x}||^2 \tag{4.61}$$

which translates to

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

Minimization



if $f(x,y)$ is convex in (x,y) and C is a convex set, then

$$g(x) = \inf_{y\in C} f(x,y)$$

is

contrast

$g(x) = \sup_{y\in A} f(x,y)$ is convex if $f(x,y)$ is convex in x for every $y \in A$

examples

- $f(x,y) = x^TAx + 2x^TBy + y^TCy$ with

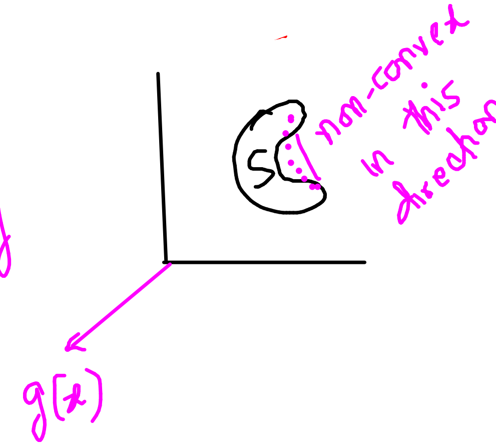
so that inf over y exists

$$\underbrace{\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0}_{\text{convexity of } f(x,y)}, \qquad C \succ 0$$

  minimizing over y gives $g(x) = \inf_y f(x,y) =$

  g is convex, hence

- distance to a set: $\mathbf{dist}(x,S) = \inf_{y\in S} \|x-y\|$ is



Next result: $g(x)$ is convex in $\mathbb{R}^2$ iff its projection along every line in $\mathbb{R}^2$ is convex

Perspective

the **perspective** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$,

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

g is convex if

examples

Minimization

if $f(x,y)$ is convex in (x,y) and C is a convex set, then

$$g(x) = \inf_{y \in C} f(x,y)$$

is convex

Gauss elimination on submatrices

Result of Gauss elimination on submatrices

examples

- $f(x,y) = x^T A x + 2x^T B y + y^T C y$ with

BC^{-1}
$$\left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right] \succeq 0, \qquad C \succ 0$$

  minimizing over y gives $g(x) = \inf_y f(x,y) = x^T(A - BC^{-1}B^T)x$

  g is convex, hence Schur complement $A - BC^{-1}B^T \succeq 0$

- distance to a set: $\mathbf{dist}(x,S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex

Perspective

the **perspective** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$,

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

g is convex if f is convex

examples

- $f(x) = x^T x$ is convex; hence $g(x,t) = x^T x/t$ is convex for $t > 0$
- negative logarithm $f(x) = -\log x$ is convex; hence relative entropy $g(x,t) = t\log t - t\log x$ is convex on $\mathbf{R}^2_{++}$
- if f is convex, then

$$g(x) = (c^T x + d) f\left((Ax+b)/(c^T x + d)\right)$$

  is convex on $\{x \mid c^T x + d > 0,\ (Ax+b)/(c^T x + d) \in \mathbf{dom}\, f\}$

} Linear fractional fn

Restriction of a convex function to a line

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$,

$$g(t) = f(x + tv), \qquad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

is convex (in t) for any $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

can check convexity of f by checking convexity of functions of one variable

example. $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, X = \mathbf{S}^n_{++}$ $(\lambda_i > 0)$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \underbrace{\log\det X}_{\text{constant}} + \log\det(I + tX^{-1}V) \\ &= \log\det X + \underbrace{\log \prod_i \left(1 + t\lambda_i(X^{-1}V)\right)}_{\text{use the fact that } \det(A) = \prod_i \lambda_i(A)} \end{aligned}$$

$[\det(AB) = \det(A)\det(B)]$

S^n

Restriction of a convex function to a line

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$,

$$g(t) = f(x + tv), \qquad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

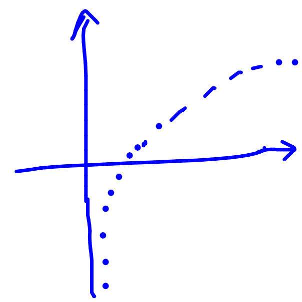
is convex (in t) for any $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

can check convexity of f by checking convexity of functions of one vari

example. $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, X = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

constant



where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$

g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

What about closedness? H/w