EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
Lecture 7

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

April 24, 2012

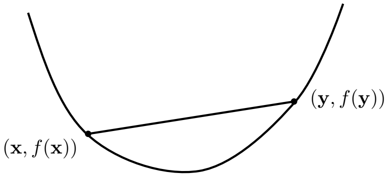
Outline

Convex Functions: Definition

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* if $\mathrm{dom}(f)$ is a convex set and

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

for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$ and $0 \leq \theta \leq 1$. Geometrically, this means that the *chord* or line segment between any two points $(\mathbf{x}, f(\mathbf{x}))$ and $(\mathbf{y}, f(\mathbf{y}))$ lies above the graph of f.



- The function f from above is said to be *strictly convex* if $\mathrm{dom}(f)$ is convex and

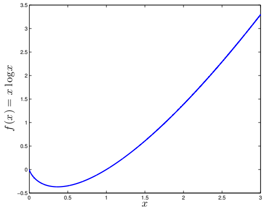
  $$f(\theta \mathbf{x} + (1-\theta)\,\mathbf{y}) < \theta f(\mathbf{x}) + (1-\theta)\, f(\mathbf{y}) \ ,$$

  for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$ such that $\mathbf{x} \neq \mathbf{y}$ and $0 < \theta < 1$.
- The function f is *concave* if $-f$ is convex. Similarly, f is *strictly concave* if $-f$ is strictly convex.
- The function f is *affine* if and only if f is both convex and concave, meaning we have equality in the top expression.

Convex/Concave Functions: Examples on $\mathbb{R}$

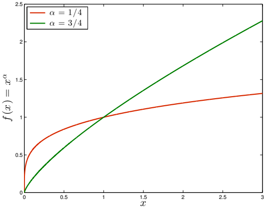
Convex Functions:

- Affine: $ax + b$ on $\mathbb{R}$, for any $a, b \in \mathbb{R}$.
- Exponential: e^{ax}, for any $a \in \mathbb{R}$.
- Powers: x^α on $\mathbb{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$.
- Powers of absolute value: $|x|^p$ on $\mathbb{R}$, for $p \geq 1$.
- Negative entropy: $x \log x$ on $\mathbb{R}_{++}$.



Concave Functions:

- Affine: $ax + b$ on $\mathbb{R}$, for any $a, b \in \mathbb{R}$.
- Logarithm: $\log x$ on $\mathbb{R}_{++}$.
- Powers: x^α on $\mathbb{R}_{++}$, for $0 \leq \alpha \leq 1$.



Convex/Concave Functions: Examples on $\mathbb{R}^n$ & $\mathbb{R}^{m \times n}$

The notion of convexity/concavity easily generalizes to functions over matrices. For example, a function $f : \mathbb{R}^{m \times n} \to \mathbb{R}$ is convex if $\mathrm{dom}(f)$ is a convex set and

$$f(\theta \mathbf{X} + (1-\theta)\,\mathbf{Y}) \leq \theta f(\mathbf{X}) + (1-\theta)\, f(\mathbf{Y}) \ ,$$

for all $\mathbf{X}, \mathbf{Y} \in \mathrm{dom}(f)$ and $0 \leq \theta \leq 1$.

All affine functions are both convex and concave. Also, all norms are convex.

Examples on $\mathbb{R}^n$:

- Affine function:

$$f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} + b \,.$$

- ℓ_p-norms:

$$||\mathbf{x}||_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} \text{ for } p \geq 1 \,; \ ||\mathbf{x}||_\infty = \max_{k=1,\ldots,n} |x_k| \ .$$

Examples on $\mathbb{R}^{m \times n}$:

- Affine function:

$$f(\mathbf{X}) = \mathrm{tr}\left(\mathbf{A}^T \mathbf{X}\right) + b = \sum_{k=1}^{m} \sum_{\ell=1}^{n} A_{k,\ell} X_{k,\ell} + b \,.$$

- Spectral (maximum singular value) norm:

$$f(\mathbf{X}) = ||\mathbf{X}||_2 = \sigma_{\max}(\mathbf{X}) = \sqrt{\lambda_{\max}(\mathbf{X}^T \mathbf{X})} \,.$$

Restriction of a Convex Function to a Line

It can be shown that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if and only if the function $g : \mathbb{R} \rightarrow \mathbb{R}$ given by

$$g(t) \triangleq f(\mathbf{x} + t\mathbf{v}) \ , \ \mathrm{dom}(g) = \{t : \mathbf{x} + t\mathbf{v} \in \mathrm{dom}(f)\} \ ,$$

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

Similarly, $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ is convex if and only if the function $g(t) \triangleq f(\mathbf{X} + t\mathbf{V})$, with $\mathrm{dom}(g) = \{t : \mathbf{X} + t\mathbf{V} \in \mathrm{dom}(f)\}$ is convex in t for any $\mathbf{X} \in \mathrm{dom}(f)$, $\mathbf{V} \in \mathbb{R}^{m \times n}$.

Hence, we can check for the convexity of f by checking the convexity of functions of one variable, namely $\{g(t)\}$.

Example:

- Consider $f : \mathbb{S}^n \rightarrow \mathbb{R}$ given by $f(\mathbf{X}) = \log(\det(\mathbf{X}))$, with $\mathrm{dom}(f) = \mathbb{S}^n_{++}$. Then, we have

$$\begin{aligned} g(t) = \log(\det(\mathbf{X} + t\mathbf{V})) &= \log(\det(\mathbf{X})) + \log\Big(\det\Big(\mathbf{I} + t \underbrace{\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}}_{\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T} \Big) \Big) \ , \\ &= \log(\det(\mathbf{X})) + \sum_{k=1}^{n} \log(1 + t\lambda_k) \ , \end{aligned}$$

where λ_k are the eigenvalues of $\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}$. As g is concave in t (for any choice of $\mathbf{X} \succ \mathbf{0}$ and $\mathbf{V} \in \mathbb{S}^n$), it follows that f is concave.

Extended-Value Extensions

The *extended-value extension* $\widetilde{f}$ of a function f is defined as

$$\widetilde{f}(\mathbf{x}) \triangleq \begin{cases} f(\mathbf{x}), & \mathbf{x} \in \mathrm{dom}(f) \\ \infty, & \mathbf{x} \notin \mathrm{dom}(f) \end{cases} .$$

This function $\widetilde{f}$ *extends* f to be defined on all $\mathbb{R}^n$ (or $\mathbb{R}^{m \times n}$ when applicable) and takes on values in $\mathbb{R} \cup \{\infty\}$.

It can often be used to simplify notation. For example, the condition

$$0 \leq \theta \leq 1 \Longrightarrow \widetilde{f}(\theta \mathbf{x} + (1-\theta)\,\mathbf{y}) \leq \theta \widetilde{f}(\mathbf{x}) + (1-\theta)\,\widetilde{f}(\mathbf{y}) \,, \tag{1}$$

when viewed as an inequality in $\mathbb{R} \cup \{\infty\}$, is equivalent to the two conditions:

- $\mathrm{dom}(f)$ is convex,
- for $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$,

$$0 \leq \theta \leq 1 \Longrightarrow f(\theta \mathbf{x} + (1-\theta)\,\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta)\, f(\mathbf{y}) \,.$$

In other words, f is convex if and only if (1) holds.

Jensen's Inequality

The basic definition of convexity can be used to extend it to the following form known as *Jensen's inequality*.

Jensen's Inequality:

Suppose $\mathbf{z}$ is a random variable such that $\mathbf{z} \in \mathrm{dom}(f)$ with probability one. Then, f is convex if and only if $\mathrm{dom}(f)$ is a convex set and

$$f(E[\mathbf{z}]) \leq E[f(\mathbf{z})] \ ,$$

for any such $\mathbf{z}$ as described above.

In other words, for any $p(\mathbf{z}) \geq 0$ on $\mathcal{S} \subseteq \mathrm{dom}(f)$ with $\int_{\mathcal{S}} p(\mathbf{z}) \, d\mathbf{z} = 1$, we have that f is convex if and only if $\mathrm{dom}(f)$ is a convex set and

$$f\left(\int_{\mathcal{S}} \mathbf{z} \, p(\mathbf{z}) \, d\mathbf{z} \right) \leq \int_{\mathcal{S}} f(\mathbf{z}) \, p(\mathbf{z}) \, d\mathbf{z} \, .$$

The basic inequality can be recovered by setting $p(\mathbf{z}) = \theta \delta(\mathbf{z} - \mathbf{x}) + (1 - \theta) \, \delta(\mathbf{z} - \mathbf{y})$, where $\delta(\mathbf{z})$ is the *Dirac delta function*. Using Jensen's inequality, we can also extend the basic inequality to include convex combinations of an arbitrary number of points:

$$f(\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k) \leq \theta_1 f(\mathbf{x}_1) + \cdots + \theta_k f(\mathbf{x}_k) \ ,$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathrm{dom}(f)$, $\theta_1, \ldots, \theta_k \geq 0$, and $\theta_1 + \cdots + \theta_k = 1$.

First-Order Condition for Convexity

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is *differentiable*, meaning that $\mathrm{dom}(f)$ is open and the gradient vector

$$\nabla f(\mathbf{x}) \triangleq \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} & \cdots & \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}^T ,$$

exists at each $\mathbf{x} \in \mathrm{dom}(f)$. Then we have the following necessary and sufficient first-order condition for convexity of f.

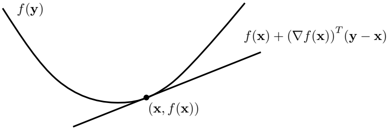
First-Order Condition for Convexity of a Differentiable Function

A differentiable function f whose domain is a convex set is convex if and only if

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

for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$.

In other words, the first-order Taylor approximation of f is a *global underestimator* of f.



Second-Order Condition for Convexity

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is *twice differentiable*, meaning that $\mathrm{dom}(f)$ is open and the Hessian matrix $\nabla^2 f(\mathbf{x}) \in \mathbb{S}^n$ given by

$$\left[\nabla^2 f(\mathbf{x})\right]_{k,\ell} \triangleq \frac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_\ell}\,,\ k,\ell = 1,\ldots,n\,,$$

exists at each $\mathbf{x} \in \mathrm{dom}(f)$. Then we have the following necessary and sufficient second-order conditions for convexity and strict convexity of f.

Second-Order Conditions for Convexity of a Twice Differentiable Function

If f is a twice differentiable function whose domain is a convex set, then

- f is convex if and only if

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}\,,\ \text{for all } \mathbf{x} \in \mathrm{dom}(f)\ .$$

- f is strictly convex if and only if

$$\nabla^2 f(\mathbf{x}) \succ \mathbf{0}\,,\ \text{for all } \mathbf{x} \in \mathrm{dom}(f)\ .$$

In other words, the function f is convex (strictly convex) if and only if the Hessian matrix $\nabla^2 f(\mathbf{x})$ is positive semidefinite (positive definite) everywhere in the domain of f.

Examples of Differentiable Convex/Concave Functions

- **Quadratic function:** $f(\mathbf{x}) = (1/2)\,\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r$ (with $\mathbf{P} \in \mathbb{S}^n$).

$$\nabla f(\mathbf{x}) = \mathbf{P}\mathbf{x} + \mathbf{q}\,,\ \nabla^2 f(\mathbf{x}) = \mathbf{P}\,.$$

  This function is convex if and only if $\mathbf{P} \succeq \mathbf{0}$.

- **Least-squares objective:** $f(\mathbf{x}) = ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2^2$.

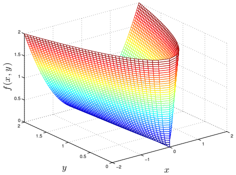
$$\nabla f(\mathbf{x}) = 2\mathbf{A}^T\left(\mathbf{A}\mathbf{x} - \mathbf{b}\right)\,,\ \nabla^2 f(\mathbf{x}) = 2\mathbf{A}^T\mathbf{A}\,.$$

  This function is convex for any $\mathbf{A}$.

- **Quadratic-over-linear:** $f(x,y) = x^2/y$.

$$\nabla f(x,y) = \frac{1}{y^2}\begin{bmatrix} 2xy \\ -x^2 \end{bmatrix}\,,\ \nabla^2 f(x,y) = \frac{2}{y^3}\begin{bmatrix} y \\ -x \end{bmatrix}\begin{bmatrix} y \\ -x \end{bmatrix}^T\,.$$

  This function is convex for $y > 0$.



Examples of Differentiable Convex/Concave Functions

- **Log-sum-exp:** The function

$$f(\mathbf{x}) = \log\left(\sum_{k=1}^{n} e^{x_k}\right) \,,$$

on $\mathbb{R}^n$ is convex. For this function, we have

$$\nabla^2 f(\mathbf{x}) = \frac{1}{\mathbf{1}^T\mathbf{z}}\operatorname{diag}(\mathbf{z}) - \frac{1}{\left(\mathbf{1}^T\mathbf{z}\right)^2}\mathbf{z}\mathbf{z}^T \,, \text{ where } z_k = e^{x_k} \,.$$

To show that $\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}$, we must verify that $\mathbf{v}^T\nabla^2 f(\mathbf{x})\,\mathbf{v} \geq 0$ for all $\mathbf{v}$, i.e.,

$$\mathbf{v}^T\nabla^2 f(\mathbf{x})\,\mathbf{v} = \frac{\left(\sum_k z_k v_k\right)\left(\sum_k z_k\right) - \left(\sum_k v_k z_k\right)^2}{\left(\sum_k z_k\right)^2} \geq 0 \,.$$

But this follows from the Cauchy-Schwarz inequality $\left(\mathbf{a}^T\mathbf{b}\right)^2 \leq ||\mathbf{a}||_2^2\,||\mathbf{b}||_2^2$, with $a_k = v_k\sqrt{z_k}$ and $b_k = \sqrt{z_k}$.

- **Geometric mean:** The function

$$f(\mathbf{x}) = \left(\prod_{k=1}^{n} x_k\right)^{1/n} \,,$$

on $\mathbb{R}_{++}^n$ is concave. The proof is similar to that for the log-sum-exp function.

Epigraph and Sublevel Sets

The link between convex functions and convex sets comes about as a result of the *epigraph* of a function. This set consists of sublevel sets of a given function.

Sublevel Sets:

The α*-sublevel set* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\mathcal{C}_\alpha \triangleq \{\mathbf{x} \in \mathrm{dom}(f) : f(\mathbf{x}) \leq \alpha\} \ .$$

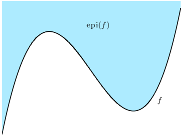
The sublevel sets of convex functions are convex, however the converse is false.

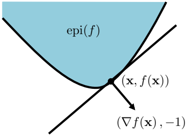
Epigraph:

The *epigraph* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\mathrm{epi}(f) \triangleq \left\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} : \mathbf{x} \in \mathrm{dom}(f) \, , f(\mathbf{x}) \leq t\right\} \ .$$

It turns out that f is a *convex function* if and only if $\mathrm{epi}(f)$ is a *convex set*.





For a differentiable convex function f, the vector $(\nabla f(\mathbf{x}) \, , -1)$ defines a *supporting hyperplane* to the epigraph of f at $\mathbf{x}$.

Practical Methods for Establishing Convexity

To establish the convexity of a given function f, the following methods are typically used.

1. Verify the definition. (This is often simplified by restricting the function to a line.)
2. For twice differentiable functions, show that the Hessian matrix is positive semidefinite, i.e., $\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}$.
3. Show that f is obtained from simple convex functions by operations that preserve convexity. (In other words, apply the *calculus of convex functions*.) Such operations include the following:
   - nonnegative weighted sum
   - composition with affine function
   - pointwise maximum and supremum
   - composition
   - minimization
   - perspective

Nonnegative Weighted Sum and Affine Composition

Nonnegative Multiple:
If f is convex, then so is αf for $\alpha \geq 0$.

Summation:
If f_1 and f_2 are convex, then so is $f_1 + f_2$. This extends to infinite sums and integrals.

Composition with Affine Function:
If f is convex, then so is $f(\mathbf{A}\mathbf{x} + \mathbf{b})$.

Examples:

- Log barrier for linear inequalities:

$$f(\mathbf{x}) = -\sum_{k=1}^{m} \log\left(b_k - \mathbf{a}_k^T \mathbf{x}\right) \ , \ \mathrm{dom}(f) = \left\{\mathbf{x} : \mathbf{a}_k^T \mathbf{x} < b_k \ , \ k = 1, \ldots, m\right\} \ .$$

- Norm of an affine function:

$$f(\mathbf{x}) = ||\mathbf{A}\mathbf{x} + \mathbf{b}|| \ .$$

  Note that this is valid for *any* norm.

Pointwise Maximum

If $f_1, \ldots, f_m$ are convex functions, then

$$f(\mathbf{x}) = \max\left\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\right\}$$

is convex.

Examples:

- Piecewise-linear function:

  $$f(\mathbf{x}) = \max_{k=1,\ldots,m}\left(\mathbf{a}_k^T\mathbf{x} + b_k\right)$$

  is convex.

- Sum of r largest components of $\mathbf{x} \in \mathbb{R}^n$:

  $$f(\mathbf{x}) = x_{[1]} + \cdots + x_{[r]}$$

  is convex, where $x_{[k]}$ denotes the k-th largest component of $\mathbf{x}$.
  The proof of this can be seen by expressing f as

  $$f(\mathbf{x}) = \max\left\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} : 1 \leq i_r < i_2 < \cdots < i_r \leq n\right\} .$$

  In other words, $f(\mathbf{x})$ is the maximum of all possible sums of r different components of $\mathbf{x}$. As f is the pointwise maximum of $C(n,r) = n!/\left(r!\,(n-r)!\right)$ linear functions, it is convex.

Pointwise Supremum

If $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$, then

$$g(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y})$$

is convex. Note that $\mathcal{A}$ *need not be a convex set*.

Examples:

- Support function of a set $\mathcal{C}$:

$$S_{\mathcal{C}}(\mathbf{x}) = \sup \left\{ \mathbf{x}^T \mathbf{y} : \mathbf{y} \in \mathcal{C} \right\}$$

  is convex.

- Distance to the farthest point in a set $\mathcal{C}$:

$$f(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{C}} ||\mathbf{x} - \mathbf{y}||$$

  is convex for any norm $||\cdot||$.

- Maximum eigenvalue of a symmetric matrix:

  For $\mathbf{X} \in \mathbb{S}^n$, $\lambda_{\max}(\mathbf{X}) = \sup_{||\mathbf{y}||_2 = 1} \mathbf{y}^T \mathbf{X} \mathbf{y}$

  is convex.

Composition with Scalar Functions

The composition $f = h \circ g$ of $g : \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R} \to \mathbb{R}$ is given by

$$f(\mathbf{x}) = h(g(\mathbf{x})) \ .$$

We have the following:

$$f \text{ is convex if } \begin{cases} g \text{ is convex, } h \text{ is convex, and } \widetilde{h} \text{ is nondecreasing} \\ g \text{ is concave, } h \text{ is convex, and } \widetilde{h} \text{ is nonincreasing} \end{cases} .$$

Here, $\widetilde{h}$ is the *extended-value extension* of h. Note that monotonicity of $\widetilde{h}$ must hold on all of $\mathbb{R}$.

Proof: (for $n = 1$ and differentiable g, h)
Note that

$$f''(x) = h''(g(x)) \left(g'(x)\right)^2 + h'(g(x))\, g''(x) \ .$$

From this, the conditions for the convexity of f are clear.

Examples:

- $e^{g(x)}$ is convex if g is convex.
- $1/g(x)$ is convex if g is concave and positive.

Vector Composition

The composition $f = h \circ \mathbf{g}$ of $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^k$ and $h : \mathbb{R}^k \to \mathbb{R}$ is given by

$$f(\mathbf{x}) = h(\mathbf{g}(\mathbf{x})) = h(g_1(\mathbf{x}), \ldots, g_k(\mathbf{x})) \ .$$

We have the following:

$$f \text{ is convex if } \begin{cases} g_i \text{ is convex, } h \text{ is convex, and } \widetilde{h} \text{ is nondecreasing in each argument} \\ g_i \text{ is concave, } h \text{ is convex, and } \widetilde{h} \text{ is nonincreasing in each argument} \end{cases} .$$

As before, $\widetilde{h}$ is the *extended-value extension* of h. Note that monotonicity of $\widetilde{h}$ must hold in each argument on all of $\mathbb{R}$.

Proof: (for $n = 1$ and differentiable $\mathbf{g}, h$)
Note that

$$f''(x) = \left(\mathbf{g}'(x)\right)^T \left(\nabla^2 h(\mathbf{g}(x))\right) \left(\mathbf{g}'(x)\right) + \left(\nabla h(\mathbf{g}(x))\right)^T \left(\mathbf{g}''(x)\right) \ .$$

From this, the conditions for the convexity of f are clear.

Examples:

- $\sum_{i=1}^m \log g_i(\mathbf{x})$ is concave if g_i is concave and positive for every i.
- $\log \left(\sum_{i=1}^m \exp g_i(\mathbf{x})\right)$ is convex if g_i is convex for every i.

Minimization

If $f(\mathbf{x}, \mathbf{y})$ is convex in $(\mathbf{x}, \mathbf{y})$ and $\mathcal{C}$ is a convex set, then

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

is convex. Note that $\mathcal{C}$ *need be a convex set*.

Examples:

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

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \succeq \mathbf{0} \, .$$

  Minimizing over $\mathbf{y}$ gives

$$g(\mathbf{x}) = \inf_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \left(\mathbf{A} - \mathbf{B} \mathbf{C}^{\#} \mathbf{B}^T \right) \mathbf{x} \, .$$

  Since g is convex, the Schur complement is positive semidefinite, i.e., $\left(\mathbf{A} - \mathbf{B} \mathbf{C}^{\#} \mathbf{B}^T \right) \succeq \mathbf{0}$.

- Distance to a set: The distance between a point $\mathbf{x}$ and set $\mathcal{S}$ is defined as

$$\operatorname{dist}(\mathbf{x}, \mathcal{S}) \triangleq \inf_{\mathbf{y} \in \mathcal{S}} ||\mathbf{x} - \mathbf{y}|| \, ,$$

  and is convex if $\mathcal{S}$ is a convex set.

Perspective

The *perspective* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the function $g : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ given by

$$g(\mathbf{x}, t) \triangleq t f(\mathbf{x}/t) \ , \ \mathrm{dom}(g) = \{(\mathbf{x}, t) : \mathrm{x}/t \in \mathrm{dom}(f) \ , \ t > 0\} \ .$$

It can be shown that g is convex if f is convex.

Examples:

- Perspective of Euclidean norm squared:
  The function $f(\mathbf{x}) = \mathbf{x}^T \mathbf{x}$ is convex, and so $g(\mathbf{x}, t) = \mathbf{x}^T \mathbf{x}/t$ is convex for $t > 0$.
- Relative entropy:
  The *negative logarithm* $f(x) = -\log x$ is convex. Thus, the *relative entropy* function $g(x, t) = t \log t - t \log x$ is convex on $\mathbb{R}^2_{++}$. This function plays a role in the *Kullback-Leibler divergence* between $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n_{++}$, given by

  $$D_{\mathrm{kl}}(\mathbf{u}, \mathbf{v}) \triangleq \sum_{k=1}^{n} \left(u_k \log(u_k/v_k) - u_k + v_k \right) \ .$$

  It is convex, since it is the relative entropy plus a linear function of $(\mathbf{u}, \mathbf{v})$.
- Affine function of perspective of affine function:
  If f is convex, then

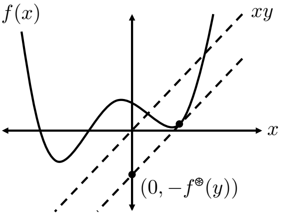
  $$g(\mathbf{x}) = \left(\mathbf{c}^T \mathbf{x} + d\right) f\left((\mathbf{A}\mathbf{x} + \mathbf{b}) \, / \left(\mathbf{c}^T \mathbf{x} + d\right) \right)$$

  is convex on $\left\{\mathbf{x} : \mathbf{c}^T \mathbf{x} + d > 0 \, , \ (\mathbf{A}\mathbf{x} + \mathbf{b}) \, / \left(\mathbf{c}^T \mathbf{x} + d\right) \in \mathrm{dom}(f)\right\}$.

Convex Conjugate: Definition and Properties

The *convex conjugate* or *Legendre-Fenchel transformation* of a function f is denoted $f^{\circledast}$ and defined as

$$f^{\circledast}(\mathbf{y}) \triangleq \sup_{\mathbf{x}\in\mathrm{dom}(f)} \left\{ \mathrm{Re}[\langle \mathbf{x}, \mathbf{y} \rangle] - f(\mathbf{x}) \right\} .$$



Properties:

- The conjugate $f^{\circledast}$ is always convex, even if f is not.
- The conjugate of the conjugate always satisfies $f^{\circledast\circledast} \leq f$ for any f.
- The conjugate of the conjugate satisfies $f^{\circledast\circledast} = f$ if and only if f is a *proper*, *lower semi-continuous*, *convex* function or f is identically $\pm\infty$.
- The conjugate $f^{\circledast}$ will play a role in *duality* when we come to the *Lagrange dual problem*.

Convex Conjugate: Examples

- Negative logarithm: $f(x) = -\log x$.

$$f^{\circledast}(y) = \sup_{x>0} \{xy + \log x\} \,, \\ = \begin{cases} -1 - \log(-y)\,, & y < 0 \\ \infty\,, & y \geq 0 \end{cases} .$$

- Strictly convex quadratic: $f(\mathbf{x}) = (1/2)\,\mathbf{x}^T\mathbf{Q}\mathbf{x}$ with $\mathbf{Q} \in \mathbb{S}^n_{++}$.

$$f^{\circledast}(\mathbf{y}) = \sup_{\mathbf{x}} \left\{\mathbf{y}^T\mathbf{x} - (1/2)\,\mathbf{x}^T\mathbf{Q}\mathbf{x}\right\} = (1/2)\,\mathbf{y}^T\mathbf{Q}^{-1}\mathbf{y}\,.$$

- Log-determinant: $f(\mathbf{X}) = \log\det\mathbf{X}^{-1}$ on $\mathbb{S}^n_{++}$.

$$f^{\circledast}(\mathbf{Y}) = \sup_{\mathbf{X}\succ\mathbf{0}} \{\mathrm{tr}(\mathbf{Y}\mathbf{X}) + \log\det\mathbf{X}\} \,, \\ = \log\det\left((-\mathbf{Y})^{-1}\right) - n\,, \text{ with } \mathrm{dom}\left(f^{\circledast}\right) = -\mathbb{S}^n_{++}\,.$$

- Indicator function: The *indicator function* $I_{\mathcal{C}}(\mathbf{x})$ of a set $\mathcal{C}$ is defined to be 0 if $\mathbf{x} \in \mathcal{C}$ and ∞ if $\mathbf{x} \notin \mathcal{C}$. Its convex conjugate is given by

$$I_{\mathcal{C}}^{\circledast}(\mathbf{y}) = \sup_{\mathbf{x}\in\mathcal{C}} \mathbf{y}^T\mathbf{x}\,,$$

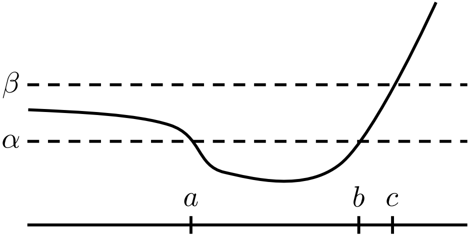
  which is the *support function* of the set $\mathcal{C}$.

Quasiconvex Functions: Definition

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *quasiconvex* if $\text{dom}(f)$ is convex and the sublevel sets

$$\mathcal{S}_\alpha \triangleq \{\mathbf{x} \in \text{dom}(f) : f(\mathbf{x}) \leq \alpha\}$$

are convex for all α.



- f is said to be quasiconcave if $-f$ is quasiconvex.
- f is said to be quasilinear if it is quasiconvex and quasiconcave.

Quasiconvex Functions: Examples

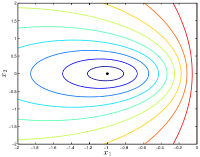
- $\sqrt{|x|}$ is quasiconvex on $\mathbb{R}$.
- $\text{ceil}(x) = \lceil x \rceil = \inf \{z \in \mathbb{Z} : z \geq x\}$ and $\text{floor}(x) = \lfloor x \rfloor = \sup \{z \in \mathbb{Z} : z \leq x\}$ are quasilinear.
- $\log x$ is quasilinear on $\mathbb{R}_{++}$.
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbb{R}^2_{++}$.
- The linear-fractional function

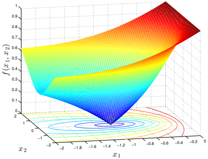
$$f(\mathbf{x}) = \frac{\mathbf{a}^T \mathbf{x} + b}{\mathbf{c}^T \mathbf{x} + d}, \ \text{dom}(f) = \left\{ \mathbf{x} : \mathbf{c}^T \mathbf{x} + d > 0 \right\},$$

  is quasilinear.
- The *distance ratio function* given by

$$f(\mathbf{x}) = \frac{||\mathbf{x} - \mathbf{a}||_2}{||\mathbf{x} - \mathbf{b}||_2}, \ \text{dom}(f) = \left\{ \mathbf{x} : ||\mathbf{x} - \mathbf{a}||_2 \leq ||\mathbf{x} - \mathbf{b}||_2 \right\},$$

  is quasiconvex.





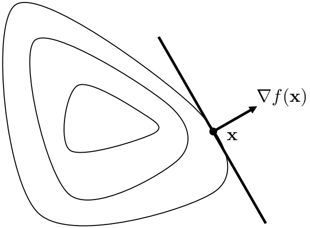
Quasiconvex Functions: Various Properties

Modified Jensen's Inequality: A function f is quasiconvex if and only if $\mathrm{dom}(f)$ is convex and for any $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$ and $0 \leq \theta \leq 1$, we have

$$f(\theta \mathbf{x} + (1-\theta)\,\mathbf{y}) \leq \max\left\{f(\mathbf{x})\,,\, f(\mathbf{y})\right\}\;.$$

First-Order Condition: If f is differentiable, then f is quasiconvex if and only if $\mathrm{dom}(f)$ is convex and for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$, we have

$$f(\mathbf{y}) \leq f(\mathbf{x}) \Longrightarrow \left(\nabla f(\mathbf{x})\right)^T (\mathbf{y} - \mathbf{x}) \leq 0\,.$$



Second-Order Condition: If f is quasiconvex, then for all $\mathbf{x} \in \mathrm{dom}(f)$ and all $\mathbf{y} \in \mathbb{R}^n$, we have

$$\mathbf{y}^T \left(\nabla f(\mathbf{x})\right) = 0 \Longrightarrow \mathbf{y}^T \left(\nabla^2 f(\mathbf{x})\right) \mathbf{y} \geq 0\,.$$

Operations that Preserve Quasiconvexity

- **Nonnegative Weighted Maximum:** If $w_k \geq 0$ and f_k are quasiconvex functions for $k = 1, \ldots, m$, then
$$f = \max\left\{w_1 f_1, \ldots, w_m f_m\right\}$$
is quasiconvex. This also extends to the general pointwise supremum
$$f(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{A}} \left\{w(\mathbf{y})\, g(\mathbf{x}, \mathbf{y})\right\} \,,$$
where $w(\mathbf{y}) \geq 0$ and $g(\mathbf{x}, \mathbf{y})$ is quasiconvex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$.
- **Composition:** If $g : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex and $h : \mathbb{R} \to \mathbb{R}$ is nondecreasing, then $f = h \circ g = h(g(\mathbf{x}))$ is quasiconvex.
- **Minimization:** If $f(\mathbf{x}, \mathbf{y})$ is jointly quasiconvex in $\mathbf{x}$ and $\mathbf{y}$ and $\mathcal{C}$ is a convex set, then
$$g(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{C}} f(\mathbf{x}, \mathbf{y}) \,,$$
is quasiconvex.

In general, the conditions guaranteeing quasiconvexity are weaker than those guaranteeing convexity. For example, *sums* of quasiconvex functions are not necessarily quasiconvex.

Log-Concave/Convex Functions

A positive function f is said to be *log-concave* if $\log f$ is concave. Equivalently, such a function is log-concave if and only if

$$f(\theta \mathbf{x} + (1-\theta)\,\mathbf{y}) \geq f(\mathbf{x})^{\theta}\, f(\mathbf{y})^{(1-\theta)} \text{ for } 0 \leq \theta \leq 1 \, .$$

Analogously, a function f is said to be *log-convex* if $\log f$ is convex.

Examples:

- Powers: x^a on $\mathbb{R}_{++}$ is log-convex for $a \leq 0$ and log-concave for $a \geq 0$.
- Common probability density functions (pdfs): Several pdfs, such as that for the normal (or Gaussian) distribution ($\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$) given by

  $$f_{\mathbf{x}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \, ,$$

  are log-concave.
- Cumulative distribution function (cdf) of the standard normal distribution: If $\mathrm{x} \sim \mathcal{N}(0,1)$, then the cdf of x given by

  $$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du \, ,$$

  is log-concave.

Fundamental Properties of Log-Concave Functions

- **Twice Differentiable Log-Concave/Convex Functions:**
  If f is twice differentiable with a convex domain, then f is log-concave/convex if and only if

$$f(\mathbf{x})\,\nabla^2 f(\mathbf{x}) \preceq \nabla f(\mathbf{x})\,\nabla f(\mathbf{x})^T \quad \text{(log-concave)}\,,$$
$$f(\mathbf{x})\,\nabla^2 f(\mathbf{x}) \succeq \nabla f(\mathbf{x})\,\nabla f(\mathbf{x})^T \quad \text{(log-convex)}\,,$$

  for all $x \in \mathrm{dom}(f)$. For the log-convex case, an equivalent condition is

$$\begin{bmatrix} \nabla^2 f(\mathbf{x}) & \nabla f(\mathbf{x}) \\ \nabla f(\mathbf{x})^T & f(\mathbf{x}) \end{bmatrix} \succeq \mathbf{0}\,,$$

  for all $x \in \mathrm{dom}(f)$, as $f(\mathbf{x}) > 0$ in this region.
- **Multiplication, Addition, and Integration:**
  - The product of log-concave functions is log-concave.
  - The sum of log-concave functions is not always log-concave.
  - If $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is log-concave, then

$$g(\mathbf{x}) = \int f(\mathbf{x}, \mathbf{y})\; d\mathbf{y}$$

    is log-concave.

Consequences of Integration Property

- The convolution $f * g$ given by

$$(f * g)(\mathbf{x}) = \int f(\mathbf{x} - \mathbf{y})\, g(\mathbf{y})\ d\mathbf{y} \,,$$

  of two log-concave functions f, g is log-concave.

- If $\mathcal{C} \subseteq \mathbb{R}^n$ is convex and $\mathbf{y}$ is a random variable with log-concave pdf, then

$$f(\mathbf{x}) \triangleq \Pr\{\mathbf{x} + \mathbf{y} \in \mathcal{C}\} \ ,$$

  is a log-concave function.
  The proof of this follows by expressing $f(\mathbf{x})$ is an integral of the product of two log-concave functions. Specifically, if $p(\mathbf{y})$ denotes the pdf of $\mathbf{y}$ and we define $g(\mathbf{u})$ as

$$g(\mathbf{u}) \triangleq \begin{cases} 1\,, & \mathbf{u} \in \mathcal{C} \\ 0\,, & \mathbf{u} \notin \mathcal{C} \end{cases} \,,$$

  then we have

$$f(\mathbf{x}) = \int g(\mathbf{x} + \mathbf{y})\, \mathbf{p}(\mathbf{y})\ d\mathbf{y} \,.$$

  As $g(\mathbf{u})$ and $p(\mathbf{y})$ are log-concave, so too is $f(\mathbf{x})$.

Yield Function Example

In a manufacturing process, the *yield function* measures the likelihood of meeting an acceptable value of yield given a nominal or target operating point. It is given by

$$Y(\mathbf{x}) = \Pr\{\mathbf{x} + \mathbf{w} \in \mathcal{S}\} \ ,$$

where we have

- $\mathbf{x} \in \mathbb{R}^n$: nominal parameter values for product,
- $\mathbf{w} \in \mathbb{R}^n$: random variations of parameters in manufactured product,
- $\mathcal{S}$: set of acceptable values.

If $\mathcal{S}$ is convex and $\mathbf{w}$ has a log-concave pdf, then

- Y is log-concave,
- yield regions $\{\mathbf{x} : Y(\mathbf{x}) \geq \alpha\}$ are convex.

Convexity with respect to a Generalized Inequality

If $\mathcal{K} \subseteq \mathbb{R}^m$ is a proper cone with an associated generalized inequality $\preceq_{\mathcal{K}}$, then we say that $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is *$\mathcal{K}$-convex* if $\mathrm{dom}(\mathbf{f})$ is convex and

$$\mathbf{f}(\theta\mathbf{x} + (1-\theta)\,\mathbf{y}) \preceq_{\mathcal{K}} \theta\mathbf{f}(\mathbf{x}) + (1-\theta)\,\mathbf{f}(\mathbf{y}) \ ,$$

for all $\mathbf{x}, \mathbf{y} \in \mathrm{dom}(\mathbf{f})$ and $0 \leq \theta \leq 1$. Similarly, we say that $\mathbf{f}$ is *strictly $\mathcal{K}$-convex* if

$$\mathbf{f}(\theta\mathbf{x} + (1-\theta)\,\mathbf{y}) \prec_{\mathcal{K}} \theta\mathbf{f}(\mathbf{x}) + (1-\theta)\,\mathbf{f}(\mathbf{y}) \ ,$$

for all $\mathbf{x} \neq \mathbf{y}$ and $0 < \theta < 1$.

Example:

Suppose $\mathbf{F} : \mathbb{R}^{m \times n} \to \mathbb{S}_+^n$ is given by $\mathbf{F}(\mathbf{X}) = \mathbf{X}^T\mathbf{X}$, where $\mathcal{K} = \mathbb{S}_+^n$. Then we can show that $\mathbf{F}$ is $\mathbb{S}_+^n$-convex.

Proof: For fixed $\mathbf{z} \in \mathbb{R}^n$, the function $f(\mathbf{X}) \triangleq \mathbf{z}^T\mathbf{X}^T\mathbf{X}\mathbf{z} = ||\mathbf{X}\mathbf{z}||_2^2$ is a convex quadratic function of the components of $\mathbf{X}$. From this, it follows that $\mathbf{F}(\mathbf{X})$ is $\mathbb{S}_+^n$-convex.

As a result, we conclude that for any $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{m \times n}$, we have

$$(\theta\mathbf{X} + (1-\theta)\,\mathbf{Y})^T\,(\theta\mathbf{X} + (1-\theta)\,\mathbf{Y}) \preceq \theta\mathbf{X}^T\mathbf{X} + (1-\theta)\,\mathbf{Y}^T\mathbf{Y} \,.$$