CONVEX SETS

Def: A set is convex is every line between two points stays in the set?

\[ \theta x_1 + (1 - \theta)x_2, \quad 0 \leq \theta \leq 1 \]

More General:

All \textbf{convex combinations} lie in set

\[
\sum \theta_i x_i = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \\
\sum \theta_i = \theta_1 + \theta_2 + \theta_3 = 1, \quad \theta_i \geq 0 \forall i
\]

These definitions are the same!
IS IT CONVEX?

What if the square was round?

figures form Boyd and Vandenberghe
CONVEX HULL

All convex combinations of points an a set

…it’s always the smallest convex superset
important examples?

Are these convex?

Hyperplane \( \{x | a^T x = b\} \)

Half-space \( \{x | a^T x \geq b\} \)

Sphere \( \{x | \|x - x_0\| = b\} \)

Ball \( \{x | \|x - x_0\| \leq b\} \)

Polynomials \( \{f | f = \sum_i a_i x_i^i\} \)
FUNCTIONS OVER NON-CONVEX DOMAIN

$\Omega$

$f : \Omega \rightarrow \mathbb{R}$
EXAMPLE:
LINEAR STRESS IN FINITE ELEMENTS

Functions space over non-convex domain is **convex**

Biharmonic equation
\[
\min_{u: \Omega \to \mathbb{R}} \| \Delta u \|^2 - \langle u, r \rangle
\]
UNIT BALL

\[ \{ x \mid \|x\| \leq b \} \]

Convex??
Does it depend on which norm??

No! Because of triangle inequality.
CONES

\[ x \in C \implies ax \in C, \quad \forall a \geq 0 \]

Second-order cone:

\[ C_2 = \{(x, t)\mid \|x\| \leq t\} \in \mathbb{R}^{n+1} \]
SEMIDEFINITE CONE

\[ S^n = \{ A \in \mathbb{R}^{n \times n} | A = A^T, A \succeq 0 \} \]

Set of PSD matrices?

Why is this a cone?
ALLOWED OPERATIONS

The intersection of convex sets is convex
SIMPLEX

\[ \{ x : Ax \leq b \} \]

Is it convex? Why?
SEMIDEFINITE CONE

\[ x^T Ax \geq 0, \forall x \]

\[ \mathcal{S}_+ = \bigcap_{x \in \mathbb{R}^N} \{ A \mid x^T Ax \geq 0 \} \]

Convex???
OTHER ALLOWED OPERATIONS

Set sum

\[ A + B = \{x + y | x \in A, y \in Y\} \]

Set Product

\[ A \times B = \{(x, y) | x \in A, y \in Y\} \]

What about union?

\[ A \cup B \]
CONVEX FUNCTIONS

Jensen’s Inequality

\[ f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y) \]
EPIGRAPH

$\text{epi}(f) = \{ (x, y) | f(x) \leq y \}$
EPIGRAPH

\[ \text{epi}(f) = \{(x, y) | f(x) \leq y\} \]

convex function = convex epigraph

convex

non-convex
**SOME DEFINITIONS**

**Theorem**

“Any **proper, closed**, function with **bounded level sets** has a minimizer”

**Proper**: epigraph is non-empty

**Coercive**:

\[ \|x\| \to \infty \implies f(x) \to \infty \]

**Bounded level sets**:

\[ \forall \alpha \exists r, \ f(x) \leq \alpha \implies \|x\| \leq r \]

**Graphs**

- Bounded level sets
  
  \[ f(x) = \ln(x) \]

- Unbounded level sets
SOME DEFINITIONS

**Closed**: epigraph is closed or level sets are closed

**Lower semi-continuous**: 
\[ \forall \epsilon > 0, \exists \delta > 0, |x - x_0| < \delta \]
\[ \implies f(x) \geq f(x_0) - \epsilon \]
Proper + closed = LSC

\[ \text{epi}(f) = \{(x, y) \mid f(x) \leq y\} \]
WHY DO WE CARE ABOUT CONVEX FUNCTION?

Any closed function with bounded level sets has a minimizer (by compactness)

...but, for a convex function,

Any minimizer is **global** (why)
Set of minima is **convex** (why)

Therefore, we can find global minimizers
WHY DO WE CARE ABOUT CONVEX FUNCTIONS?

This can’t happen!

Minimizers of convex problems form a convex set

If you found one you found them all!
CONVEX FUNCTIONS HAVE CONVEX SUB-LEVEL SETS

\[ f(x) \]

**sub-level set**

\[ f_\alpha = \{ x | f(x) \leq \alpha \} \]
Non-convex problems can still have nice properties…
convex contours = quasi-convex < convex
strict minima and always global mins

\[ f(x) = \log(|x| + 0.1) \]

This function is “\text{log-convex}”
WHY DO WE CARE ABOUT CONVEX FUNCTION?

Any minimizer is **global** (why)
Set of minima is **convex** (why)

Therefore, we can find global minimizers

But why convex?

Convex functions are **closed under many operations**
POSITIVE WEIGHTED SUM

Sums of convex functions are convex

\[ g(x) = \sum_i f_i(x) \]

Example:
\[ f(x) = \|x\|^2 = \sum_i x_i^2 \]

Example:
\[ f(x) = x + x^2 + x^6 + |x| \]
AFFINE COMPOSITION

Affine composition with convex function = convex

\[ f(x) \rightarrow f(Ax + b) \]

What does this say about epigraphs?
EXAMPLES

Least squares

$$\|Ax - b\|^2 = \sum_i (A_i^T x - b)^2$$

SVM

$$\frac{1}{2} \|w\|^2 + C \sum_i h(1 - d_i^T w)$$

Logistic Regression

$$\sum_i \log(1 + \exp(-\ell_i d_i^T x))$$
POINTWISE-MAX

preserves convexity

\[ g(x) = \max_i f_i(x) \]

Absolute value

\[ |x| = \max\{x, -x\} \]

Infinity norm

\[ \|x\|_\infty = \max_i |x_i| \]

Max eigenvalue

\[ \|A\|_2 = \max_v v^T Av \]

What does this say about epigraphs?
WHY ARE THESE CONVEX?

Trace\[ f(X) = \text{trace}(A^T X) \] \textbf{Linear operator}

Distance over set\[ f(x) = \max_{y \in C} \|x - y\| \] \textbf{Max over convex}

Distance to set\[ f(x) = \min_{y \in C} \|x - y\| \] \textbf{Min of convex (special case)}

Max eigenvalue\[ f(x) = \|b + \sum_{i} A_i x_i\|_2 \] \textbf{Affine comp}

If $g(x,y)$ is convex, then minimizing for $y$ preserves convexity
WHY ARE THESE NON-CONVEX?

Neural Net

\[ y = \sigma(X_3\sigma(X_2\sigma(X_1D)))) \]

Dictionary learning

\[ f(X, Y) = \|XY - B\|_{fro} \]
DIFFERENTIAL PROPERTIES
IMPORTANT PROPERTY

Convex functions lie **ABOVE** their linear approximation
First-order conditions = optimality for convex funcs only

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

This doesn't happen!!

Any minima is a global minima
SECOND ORDER CONDITIONS

A smooth function is convex iff \( \nabla^2 f(x) \geq 0 \), \( \forall x \)

For non-convex functions, minima satisfy: \( \nabla^2 f(x^*) \geq 0 \)

remember - the Hessian is a good local model of a smooth function
STRONG CONVEXITY

What about when there’s no Hessian?

\[
f(y) \geq f(x) + (y - x)^T \nabla f(x) + \frac{m}{2} \|y - x\|^2
\]

holds for any convex \( f \) and \( \min \) curvature

When Hessian exists...

\[
f(y) \approx f(x) + (y - x)^T \nabla f(x) + \frac{1}{2} (y - x)^T \nabla^2 f(x) (y - x)
\]

\[
\geq f(x) + (y - x)^T \nabla f(x) + \frac{\lambda_{\min}}{2} \|y - x\|^2
\]

...and so \( \nabla^2 f(x) \succeq mI \)
UPPER BOUND ON CURVATURE

Lower
\[
f(y) \geq f(x) + (y - x)^T \nabla f(x) + \frac{m}{2} \|y - x\|^2
\]

upper
\[
f(y) \leq f(x) + (y - x)^T \nabla f(x) + \frac{M}{2} \|y - x\|^2
\]

\[
f(y) \approx f(x) + (y - x)^T \nabla f(x) + \frac{1}{2} (y - x)^T \nabla^2 f(x)(y - x)
\]
\[
\leq f(x) + (y - x)^T \nabla f(x) + \frac{\lambda_{max}}{2} \|y - x\|^2
\]
LIPSCHITZ CONSTANT

\[ \| \nabla f(x) - \nabla f(y) \| \leq M \| y - x \| \]

\[ f(y) \leq f(x) + (y - x)^T \nabla f(x) + \frac{M}{2} \| y - x \|^2 \]

We use this when there's no Hessian (ex: Huber function)
OBJECTIVE ERROR BOUNDS

Lower  \[ f(y) \geq f(x) + (y - x)^T \nabla f(x) + \frac{m}{2} \|y - x\|^2 \]

upper  \[ f(y) \leq f(x) + (y - x)^T \nabla f(x) + \frac{M}{2} \|y - x\|^2 \]

We can bound the objective error in terms of distance from minimizer

\[ f(y) - f(x^*) \geq \frac{m}{2} \|y - x^*\|^2 \]

\[ f(y) - f(x^*) \leq \frac{M}{2} \|y - x^*\|^2 \]
CONDITION NUMBER

For any function

\[ \kappa = \frac{\text{major axis}}{\text{minor axis}} \]

\[ \kappa \approx \text{cond}(\nabla^2 f(x)) \]

For smooth functions

\[ \kappa = \frac{M}{m} \]

For differentiable functions

\[ \kappa = 1.5 \]

\[ \kappa = 5 \]
**SUB-DIFFERENTIAL**

\[ \partial f(x) = \{ g : f(y) > f(x) + (y - x)^T g, \, \forall y \} \]

Optimality: \[ 0 \in \partial f(x^*) \]
EXAMPLE

\[ f(z) = |z| \]

\[ \partial f(0) = [-1, 1] \]
The (sub) gradient of any convex function is monotone

\[ \langle y - x, \nabla f(y) - \nabla f(x) \rangle \geq 0 \]

or

\[ \langle y - x, g_y - g_x \rangle \geq 0 \]

for any

\[ g_x \in \partial f(x), \quad \text{and} \quad g_y \in \partial f(y) \]

This generalizes the concept of PSD Hessian

\[ (y - x)^T (\nabla f(y) - \nabla f(x)) = (y - x)^T H (y - x) \geq 0 \]
CONJUGATE FUNCTION

\[ f^*(y) = \max_x y^T x - f(x) \]

Is it convex?
**EXAMPLE: QUADRATIC**

\[ f(x) = \frac{1}{2} x^T Q x \]

\[ f^*(y) = \max_x y^T x - \frac{1}{2} x^T Q x \]

\[ y - Q x^* = 0 \]

\[ x^* = Q^{-1} y \]

\[ f^*(y) = y^T Q^{-1} y - \frac{1}{2} y^T Q^{-1} QQ^{-1} y \]

\[ f^*(y) = \frac{1}{2} y^T Q^{-1} y \]
CONJUGATE OF NORM

\[ f(x) = \|x\| \]

\[ f^*(y) = \max_x y^T x - \|x\| \]

dual norm
\[ \|y\|_* \triangleq \max_x y^T x / \|x\| \]

Holder inequality
\[ y^T x \leq \|y\|_* \|x\| \]

\[ \|y\|_* \leq 1 \rightarrow f^* = 0 \]

\[ \|y\|_* > 1 \rightarrow f^* = \infty \]

\[ f^*(y) = \begin{cases} 
0, & \|y\|_* \leq 1 \\
\infty, & \text{otherwise} 
\end{cases} \]
EXAMPLES

Holder inequality for $p$-norms

$$|\langle x, y \rangle| \leq \|x\|_p \|y\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1$$

$$f(x) = \|x\|_2, \quad f^*(y) = \mathcal{X}_2(y) = \begin{cases} 0, & \|x\|_2 \leq 1 \\ \infty, & \|x\|_2 > 1 \end{cases}$$

$$f(x) = |x|, \quad f^*(y) = \mathcal{X}_\infty(y) = \begin{cases} 0, & \|x\|_\infty \leq 1 \\ \infty, & \|x\|_\infty > 1 \end{cases}$$

$$f(x) = \|x\|_\infty, \quad f^*(y) = \mathcal{X}_1(y) = \begin{cases} 0, & |x| \leq 1 \\ \infty, & |x| > 1 \end{cases}$$

Useful when we study **duality**
PROPERTIES OF CONJUGATE

\[ f^*(y) = \max_x y^T x - f(x) \]

\( x \) is the point where \( f \) has gradient \( y \) (why?)

\[ y \in \partial f(x) \]

The gradient of conjugate = the “adjoint” of the gradient (why?)

\[ y = \nabla f(x) \iff \nabla f^*(y) = x \]

…and in general

\[ y \in \partial f(x) \iff x \in \partial f^*(y) \]

also important for convergence proofs!