UNIVERSITY OF CALIFORNIA, BERKELEY

Lecture 11 EE127: Convexity

Oscar Ortega

July 16, 2021

'The Future is convex' -P.Jackel, A. Kawai

1 CONVEX-SETS

given a set of points

$$\mathscr{P} = \{x_1, ..., x_m\} \in \mathbf{R}^n$$

give the **linear hull** generated by these points is the set of all possible linear combinations of the points.

$$x = \sum_{i=1}^{n} \lambda_i x_i : \lambda \in \mathbf{R}$$

the **affine hull**, $aff\mathscr{P}$ is the set generated by taking all possible linear combinations of the points in $\mathscr{P}$ under the restriction that the coefficeients λ_i sum up to 1.

$aff\mathscr{P}$ is the smallest affine set containing $\mathscr{P}$
eg: the affine hull of two points in space is the line through the two points. If we add another point, the affine hull becomes a triangle.

A linear combination of points is said to be **convex** if the scalars of the combination are all non-negative and sum up to one.

With the restriction that the scalars sum up to one we can think of any convex combination as the weighted average of the points. The set of all possible convex combinations is knows as the **convex hull**

$$convexhull(x^{(1)},...x^{(m)}) = \left(x = \sum_{i=1}^{n} \lambda_i x^{(i)} : \lambda_i \geq 0 : \sum_{i=1}^{m} \lambda_i = 1\right)$$

if we get rid of requirement that the set of scalars are non-negative, we have our definition for the **conic hull**

$$conichull(x^{(1)},...x^{(m)}) = \left(x = \sum_{i=1}^{n} \lambda_i x^{(i)} : \sum_{i=1}^{m} \lambda_i = 1\right)$$

2 Convexity

A subset $C \subset \mathbf{R}^n$ is said to convex if it contains the line segment between any two points in it:

$$x_1, x_2 \in C : \lambda \in [0,1] \to \lambda x_1 + (1-\lambda)x_2 \in C$$

a set C is a cone if $x \in C$ implies $ax \in C$ for every $a \geq 0$. A set C is said to be a convex and it is a cone.

2.1 Operations that preserve convexity

Intersections preserve convexity:
In other words, if $C_1, ..., C_m$ are convex sets. $\to C' = \bigcap_{i=1,...m} C_i$ is also a convex set.
As an example if we find the intersection of m half-spaces, (a convex set), this forms a convex set we now as a polyhedron.
Furthermore, if we consider the second order cone.

$$\mathcal{K}_n = \{(x,t), x \in \mathbf{R}^n, t \in \mathbf{R} : \|x\|_2 \leq t\}$$

we now determine it convex by noting that for a fixed value t, each half space is the set (x,t) where the norm of x is equal to t. In 3d, this would be a circle that is shifted up t units and is of radius t.

$$\mathcal{K}_n = \bigcap_{u:\|u\|_2 \leq 1} \{(x,t), x \in \mathbf{R}^n, t \in \mathbf{R} : u^T x \leq t\}$$

Affine Transformations preserve convexity.
If a map $f : \mathbf{R}^n \to \mathbf{R}^m$ and $C \subseteq R^n$, then the image set is convex. Note that this implies projections of convex sets onto subspaces are convex.
'a potato that is convex and strethed is still a potato. a potato that is moved is still a potato that is convex.'

2.2 Supporting Hyperplane Theorom

If $C \subseteq \mathbf{R}^n$ is conves and z is on the boundary of the set C, there there exists a supporting hyperplane for C at z. What this is saying is that if you have an object that is convex like a ball, a 2d object like a straight stick will only touch the ball once.

2.3 Separating Hyperplace Theorom

If C_1, C_2 are two convex sets that do not intersect then we can divide the convex sets with a hyperplane.

3 Convex Functions

recall:

$$\mathrm{dom} f = \{x \in \mathbf{R}^n : -\infty < f(x) < \infty\}$$

Consider a function f: $\mathbf{R}^n \to \mathbf{R}$ is convex if dom f is a convex set, and for all $x, y \in domF$ and all $\lambda \in [0,1]$ it holds that $f(\lambda x + (1-\lambda)y \leq \lambda f(x) + (1-\lambda) f(y)$.
We define a function to be **concave** if the negative of the function is convex.

$$f convex \iff epi(f) isaconvexset$$

recall that epi(f) is one dimension greater than the function f.
convex functions must be $+\infty$ outside their domains. consider what would occur if the graph of -log(x) approached 17 as x approached zero, the epigraph is no longer convex.
the **sublevel set** is everything below the level set of a function. it can be verified that if f is a convex function, then a sublevel set of s is a a convex set.

$$S_\alpha = \{x \in \mathbf{R^n} : \mathbf{f(x)} \leq \alpha\}$$

The two definitions are related because the epigraph of a function is a convex set.

3.1 Operations that Preserve Convexity

if $f_i : i \in \{1, ..., m\} : \mathbf{R}^n \to \mathbf{R}$ are convex functions then a linear combination of the functions will be convex iff all the scalar weights of said combination are positive.
furthermore, if we have a convex function $f : \mathbf{R}^n \to \mathbf{R}$ and define $g(x) = f(Ax+b), A \in \mathbf{R}^{n,m}, b \in \mathbf{R}^n$ then g is convex over its domain.

3.2 First-order conditions:

if f is differentiable, then f is convex iff:

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

Key: if we consider a function that is convex and consider two points x, y in the domain of the function. the limit argument implies that the gradient approximation will always underestimate the true value of the function. We can relate this back to the supporting hyperplane. the gradient of a convex function at a point $x \in \mathbf{R}^n$ divides the whole space in two halfspaces. the portion that the gradient is greater than or equal to 0, and the portion in which the gradient is less than 0.

3.3 SECOND ORDER CONDITIONS:

A twice differentiable function f is convex iff the the hessian is positive semidefinite. Consider how in one dimension, this reduces to stating that the second derivative must always be positive.

3.4 LINE RESTRICTIONS AND THE POINTWISE MAXIMUM RULE

A function f is convex iff its restriction to any line is convex.
i.e if we know f is convex the g as defined is convex:

$$g(t) = f(x_0 + tv) : x_0 \in \mathbf{R}^n and v \in \mathbf{R}$$

Finally if we have a family of convex functions $f_1, ...f_m$ and if we define $g = max(f_1, ...f_m)$, then g as defined is also convex. this is useful when knowing a function is implicitly a maximum of several functions we also know are convex. ie: the function f: $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ as defined is convex because it can be described as the pointwise maximum of linear functions of y and t,

$$f(y,t) = \|y\|_2 - t$$
$$= f(y,t) = max_{u:\|u\|_2=1} u^T y - t$$