Convex Analysis

Daniel N. Hauser

August 25th 2022

This note summarizes some of the key concepts from the third set of lectures.

Concave Functions

A set X is *convex* if any line segment connecting any two points in the set also belongs to the set. A function $f : X \to \mathbb{R}$ is concave if it's graph on any line connecting $x, y \in X$ lies above the line segment between $f(x), f(y)$ in the range.

Definition. *We say a function $f : X \to \mathbb{R}$,with domain $X \subseteq \mathbb{R}^m$ convex is* concave *if for any $\lambda \in [0, 1]$*

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

Every critical point of a concave function is a global maximum. If f is strictly concave, then the maximum is unique.[1]

For smooth enough functions that map from $\mathbb{R}$ to $\mathbb{R}$, concavity is pretty clearly the same as (i) the second derivative is negative (i.e. the derivative of the function is decreasing), (ii) the tangent line through any point lies above the function. These properties generalize

Economic "applications" of convexity (beyond finding maxes)

1. In standard economic models, the value of information is convex, e.g. if I think there's a 50% chance that I'm sick, I'd prefer finding out for sure and then making decisions instead of only acting only on my prior.
2. Many of our canonical optimization problems lead to convex or concave solutions, e.g. expenditure or cost minimization.
3. Many common non-linear (and of course all linear problems :)) optimization problems that are solved in practice are convex/concave optimization problems. For instance, common machine learning techniques like LASSO and ridge regression.

[1] Concavity/convexity are natural economic assumptions on many objects, as they capture decreasing/increasing marginals.

Theorem 1. *Suppose $f : X \to \mathbb{R}$ is twice continuously differentiable and $X \subseteq \mathbb{R}^m$ is convex. TFAE*

1. *f is concave.*
2. *For any $x, y \in X$, $f(y) \leq f(x) + \nabla f(x) \cdot (y - x)$.*
3. *For all $x \in X$, $D^2 f(x)$ is negative semidefinite.*

A convex function admits a similar characterization. [2]

[2] Similar conditions, with the weak inequality replaced with a strict inequality, are sufficient (but not necessary) for strict concavity.

If all we care about is functions where local max $\Rightarrow$ global max, we can broaden this class.

Definition. *$f : X \to \mathbb{R}$, X convex is* quasiconcave *if for any $\lambda \in [0, 1]$*

$$f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}.$$

and quasiconvex if

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}.$$

Geometrically, quasiconvex functions are exactly the set of functions with convex upper-contour sets.[3]

[3] Quasiconcavity for differentiable functions is equivalent to $f(y) \geq f(x) \Rightarrow \nabla f(x) \cdot (y-x) \geq 0$. Note that if a function satisfies $f(y) > f(x) \Rightarrow \nabla f(x) \cdot (y-x) > 0$ then critical points are global maxes, which is true for concave but not quasiconcave functions. Quasiconcavity is almost, but not quite enough for this conclusion (our old friend $f(x) = x^3$ is an easy example).

The Separating Hyperplane Theorem

Especially in light of our contour set characterization of quasiconcavity, convex sets are going to be very important for us. More generally, in many economic models convexity is a natural and relatively reasonable structure to place on the set of possible values for endogenous variables. The geometry of a convex sets has a tight connection to hyperplanes, which are very simple geometric structures.

Theorem 2 (Separating Hyperplane Theorem)**.** *Let $A, B \subseteq \mathbb{R}^m$, A closed, convex, non-empty, and B compact, convex, non-empty. Finally suppose $A \cap B = \emptyset$. Then there exists a vector $p \in \mathbb{R}^m$ and a scalar $d \in \mathbb{R}$ s.t. for all $a \in A$, $b \in B$, $p \cdot a < d < p \cdot b$.*

We can always draw a plane so that A and B lie on opposite sides of the plane $\{x : p \cdot x = d\}$. The vector p is orthogonal to the plane, determining the orientation, while d pins down the position. [4]

[4] The hyperplane has dimension $m-1$. In $\mathbb{R}^2$ for instance, a hyperplane that separates

$$A = \{(x,y) \in \mathbb{R}^2 : ||(x,y)|| \leq 1\}$$

and

$$B = \{(x,y) \in \mathbb{R}^2 : ||(x,y) - (2,2)|| \leq 1\}$$

is given by $p = (1,1)$, $d = 2$ — the points that solve the equation $x + y = 2$.