Convexity

Didier Henrion[1,2]

Draft of June 10, 2013

To appear in **The Princeton Companion to Applied Mathematics** edited by Nicholas J. Higham, Mark Dennis, Paul Glendinning, Paul Martin, Fadil Santosa and Jared Tanner and published by Princeton University Press.

The notion of convexity is central in applied mathematics. It is also used in everyday life in connection with curvature properties of a surface. For example an optical lens is said to be convex if it is bulging outwards.

Convexity appears in ancient Greek geometry, for example in the description of the five regular convex space polyhedra (platonic solids). Archimedes (ca. 250BC) seems to have been the first to give a rigorous definition of convexity, similar to the geometric definition we use today: a set is convex if it contains all line segments between each of its points.

In his study of singularities of real algebraic curves, Newton (ca. 1720) introduced a convex polygon in the plane built from the exponents of the monomials of the polynomial defining the curve; this is known as the Newton polygon. Cauchy (ca. 1840) studied convex curves and remarked, for example, that if a closed convex curve is contained in a circle, then its perimeter is smaller than that of the circle. Convex polyhedra were studied by Fourier (ca. 1825) in connection with the problem of solvability of linear inequalities.

A central figure in the modern development of convexity is H. Minkowski, who was motivated by problems from number theory. In 1891, Minkowski proved that, in Euclidean space $\mathbb{R}^n$, every compact convex set with center at the origin and volume greater than 2^n contains at least one point with integer coordinates different from the origin. From Minkowski's work follows the classical isoperimetric inequality stating that among all convex sets with given volume, the ball is the one with minimal surface area. In 1896, Minkowski considered systems of the form $Ax \geq 0$ where A is a real $m \times n$ matrix and $x \in \mathbb{R}^n$. Together with the above-mentioned contribution by Fourier, this set the ground for *linear programming*, which emerged in the late 1940s, with key contributions by L.

[1]LAAS-CNRS, University of Toulouse, France.
[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic.

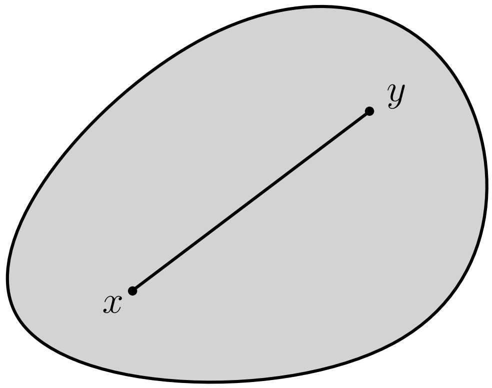

Figure 1: A convex set.

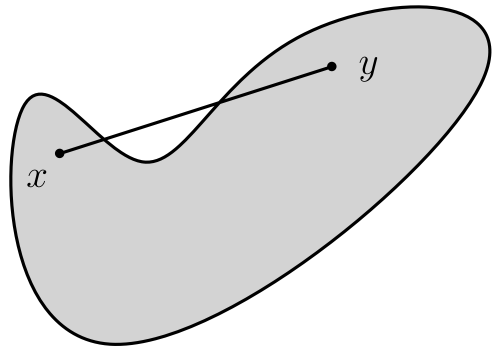

Figure 2: A nonconvex set.

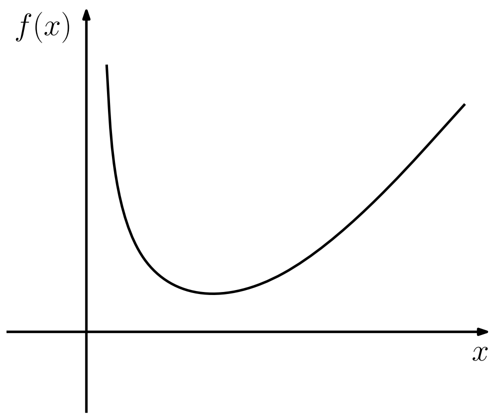

Figure 3: A convex function.

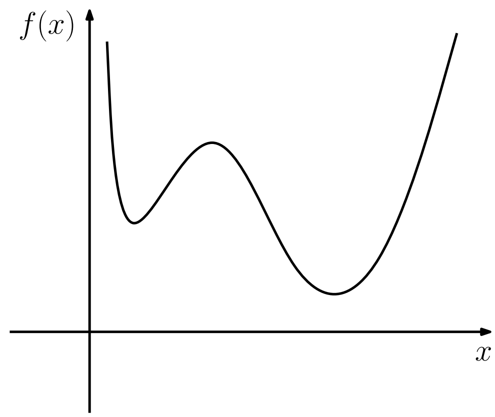

Figure 4: A nonconvex function.

Kantorovich (1912–1986) and G. Dantzig (1914–2005). In the second half of the 20th century, convexity was developed further under the impetus of W. Fenchel (1905–1988), J. J. Moreau (1923–) and R. T. Rockafellar (1935–), among many others. Convexity is now a key notion in many branches of applied mathematics: it is essential in mathematical programming (to ensure convergence of optimization algorithms), functional analysis (to ensure existence and uniqueness of solutions of problems of calculus of variations and optimal control), geometry (to classify sets and their invariants or to relate geometrical quantities), and probability and statistics (to derive inequalities).

Geometrically speaking, convex objects can be thought of as the opposite in some sense to fractal objects. Indeed, fractal objects arise in maximization problems (sponges, lungs, batteries) and they have a rough boundary. In contrast, convex objects arise in minimization problems (isoperimetric problems, smallest energy) and they have a smoother boundary.

Mathematically, a set X is *convex* if, for all $x, y \in X$, and for all $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in X$; see figures 1 and 2. Geometrically, this means that the line segment between any two points of the set belongs to the set. A real-valued function $f : X \to \mathbb{R}$ is *convex* if, for all $x, y \in X$, and for all $\lambda \in [0,1]$, it holds $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$.

Geometrically, this means that the line segment between any two points on the graph of the function lies above the graph; see figures 3 and 4. This is the same as saying that the *epigraph* $\{(x, y) : x \in X,\ y \geq f(x)\}$ is a convex set. If a function is twice continuously differentiable, convexity of the function is equivalent to nonnegativity of the quadratic form of the matrix of second-order partial derivatives (the Hessian).

If the function f is convex then $f(\lambda_1 x_1 + \cdots + \lambda_m x_m) \leq \lambda_1 f(x_1) + \cdots + \lambda_m f(x_m)$ for all $x_1, \ldots, x_m$ in X and λ in the m-dimensional unit simplex $\{\lambda \in \mathbb{R}^m : \lambda_1 + \cdots + \lambda_m = 1,\ \lambda_1 \geq 0, \ldots, \lambda_m \geq 0\}$. This is called Jensen's inequality, and more generally it can be expressed as $f(\int x\mu(dx)) \leq \int f(x)\mu(dx)$ for every probability measure μ supported on X, or equivalently, as $f(E[x]) \leq E[f(x)]$ where E denotes the expectation of a random variable.

A function f is *concave* whenever the function $-f$ is convex. If a function f is both convex and concave, it is affine. For this reason, convexity can be sometimes interpreted as a one-sided linearity, and in some instances (for example, in problems of calculus of variations and partial differential equations), nonlinear convex functions behave similarly to linear functions.

A set X is a *cone* if $x \in X$ implies $\lambda x \in X$ for all $\lambda \geq 0$. A convex cone is therefore a set that is closed under addition and under multiplication by positive scalars. Convex cones are central in optimization, and conic programming is the minimization of a linear function over an affine section of a convex cone. Important examples of convex cones include the linear cone (also called the positive orthant), the quadratic cone (also called the Lorentz cone), and the semidefinite cone (which is the set of non-negative quadratic forms, or equivalently, the set of positive semidefinite matrices).

The *convex hull* of a set X is the smallest closed convex set containing X, sometimes denoted conv X. If X is the union of a finite number of points then conv X is the polytope with vertices among these points. A theorem by Carathéodory states that given a set $X \subset \mathbb{R}^{n-1}$, every point of conv X can be expressed as $\lambda_1 x_1 + \cdots + \lambda_n x_n$ for some choice of $x_1, \ldots, x_n$ in X and λ in the n-dimensional unit simplex.

A theorem of Minkowski (generalized to infinite-dimensional spaces in 1940 by Krein and Milman) states that every compact convex set is the closure of the convex hull of its extreme points (a point $x \in X$ is *extreme* if $x = \frac{x_1 + x_2}{2}$ for some $x_1, x_2 \in X$ implies $x_1 = x_2$). Finally, we mention the Brunn–Minkowski theorem which relates the volume of the sum of two compact convex sets (all points that can be obtained by adding a point of the first set to a point of the second set) to the respective volumes of the sets.

Acknowledgments

I am grateful to Nicolas J. Higham and Jean-Baptiste Hiriart-Urruty for their comments and suggestions that significantly improved my original text. I would also like to thank Sam Clark and Princeton University Press for allowing me to distribute this preprint.

References

[1] A. Ben-Tal, A. Nemirovski. Lectures on modern convex optimization. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.

[2] M. Berger. Geometry revealed: a Jacob's ladder to modern higher geometry. Springer-Verlag, Berlin, 2010.

[3] P. M. Gruber, History of convexity, in: P. M. Gruber, J. M. Wills (editors), Handbook of convex geometry. Vol. A, pp. 1-15, North-Holland, Amsterdam, 1993.

[4] J. B. Hiriart-Urruty and C. Lemaréchal. Fundamentals of convex analysis. Springer-Verlag, Berlin, 2004.

[5] L. C. Evans. Partial differential equations. 2nd edition, Amer. Math. Soc., Providence, RI, 2010.