Contents

1 Basic concepts

1.1 Affine sets

Definition: Affine transformation: A single-valued mapping $T : x \to Tx$ from $\mathbf{R}^n$to $\mathbf{R}^m$is called an affine transformation if

$$T((1-\lambda)x + \lambda y) = (1-\lambda)Tx + \lambda Ty$$

for every x and y in $\mathbf{R}^n$and $\lambda \in \mathbf{R}$

---

*From http://eom.springer.de/M/m065240.htm : A mapping $\Gamma : X \to Y$ associating with each element x of a set X a subset $\Gamma(x)$ of a set Y. If for each x the set $\Gamma(x)$ consists of one element, then the mapping is called **single-valued**. A multi-valued mapping Γ can be treated as a single-valued mapping of X into 2^Y , that is, into the set of all subsets of Y.*

Theorem 1.5: (Affine transformations) The affine transformation from $\mathbf{R}^n$to $\mathbf{R}^m$are the mappings T of the form $Tx = Ax + a$, where A is a linear transformation and $a \in \mathbf{R}^n$.
Proof: *If Part:* Given that T is affine transform show that it can be written as $Ax + a$ where A is a linear transformation.
If T is affine, let $a = T0$. $Ax = Tx - a$. Then A is affine transformation with $A0 = 0$. We need to show that affine transformation such that $A0 = 0$ is linear.
A mapping is linear if it satisfies :
(a) Additivity : $A(x + y) = Ax + Ay$
(b) Homogenity of degree 1 : $A(\alpha x) = \alpha Ax$
Here we said A is affine and $A0 = 0$, we need to show that it is linear.
(a) Homogenity property:

$$\begin{aligned} A(\alpha x) &= A((1-\alpha)0 + \alpha x) \\ &= (1-\alpha)A0 + \alpha Ax \\ &= \alpha A(x) \end{aligned} \qquad (-1)$$

(b) Additivity property:

$$\begin{aligned} A(x+y) &= A(2((1-\frac{1}{2})x + \frac{1}{2}y)) \\ &= 2A(((1-\frac{1}{2})x + \frac{1}{2}y)) \\ &= 2A((1-\frac{1}{2})x) + 2A(\frac{1}{2}y) \\ &= Ax + Ay \end{aligned}$$

Only if part: Given that A is linear transformation, show that $Ax + a$ is affine transformation.

$$\begin{aligned} T((1-\lambda)x + \lambda y) &= A((1-\lambda)x + \lambda y) + a \\ &= (1-\lambda)Ax + \lambda Ay + a \\ &= (1-\lambda)Tx + \lambda Ty + a \end{aligned} \qquad (-7)$$

Theorem 4.5 Let f be a twice continuously differentiable real-valued function on an open convex set C in $\mathbf{R}^n$. The f is convex on C if and only if its Hessian $Q_x = (q_{ij}(x))$, $q_{ij}(x) = \frac{\partial^2 f}{\partial\zeta_i \partial\zeta_j}$ is positive semi-definite for every $x \in C$.

Proof: The convexity of f on C is equivalent to the convexity of the restriction of f to each line segment in C. This is the same as the convexity of the function $g(\lambda) = f(y + \lambda z)$ on the open real interval $\{\lambda | y + \lambda z \in C\}$ for each $y \in C$ and $z \in \mathbf{R}^n$. Let $x = y + \lambda z$.

$$
\begin{aligned}
g''(\lambda) &= \frac{d^2 g}{d\lambda^2} \\
g'(\lambda) &= \frac{df}{dx}\frac{dx}{d\lambda} \\
&= \begin{bmatrix} \frac{\partial f}{\partial x_1}\frac{dx_1}{d\lambda} \\ \frac{\partial f}{\partial x_2}\frac{dx_2}{d\lambda} \\ \cdot \\ \frac{\partial f}{\partial x_n}\frac{dx_n}{d\lambda} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial f}{\partial x_1} z_1 \\ \frac{\partial f}{\partial x_2} z_2 \\ \cdot \\ \frac{\partial f}{\partial x_n} z_n \end{bmatrix} \\
g''(\lambda) &= \langle z, Q_x z \rangle
\end{aligned}
$$

And from the previous theorem we know that g is convex for each $y \in C$ and $z \in \mathbf{R}^n$ if and only if $\langle z, Q_x z \rangle \geq 0$. □

Note: The rate of change of g with respect to λ for each y and z depends on the rate of change of the function along z and how much z changes with respect to λ.

A *quadratic* function

$$f(x) = \frac{1}{2}\langle x, Qx \rangle + \langle x, a \rangle + \alpha \tag{-11}$$

where Q is a symmetric $n \times n$ matrix, is convex on $\mathbf{R}^n$if and only if Q is positive semi-definite. We know that $\langle x, Qx \rangle$ is convex if Q is positive semi-definite, $\langle x, a \rangle + \alpha$ is also convex because it is affine transformation and we also have sum of convex functions are also convex.

One of the important consequences of the above theorem is that the *Euclidean norm*, $\langle x, x \rangle^{\frac{1}{2}}$, where Q is the identity matrix, is convex.

Theorem 4.6: (Level sets of convex functions are convex) For any convex function f and any $\alpha \in [-\infty, +\infty]$, the levels sets $\{x | f(x) < \alpha\}$ and $\{x | f(x) \leq \alpha\}$ are convex.

Proof: Suppose that level sets are not convex, that is, there exists α such that the set $C = \{x | f(x) < \alpha\}$ is not convex. Let the dom$f = S$. We know that from the definition of convex functions, that is, epi$f$ is convex, that S is convex. The above statement means that there exists $z \in S$, that is, $z = (1-\lambda)x + \lambda y$ where x, $y \in C$, but $z \notin C$, that mean $f(z) \geq \alpha$. By convexity of f, we have,

$$
\begin{aligned}
f((1-\lambda)x + \lambda y) &\leq (1-\lambda)f(x) + \lambda f(y) \\
&\leq \alpha
\end{aligned}
$$

But then the above statement implies that $(1-\lambda)x+\lambda y) \in C$. But $(1-\lambda)x+\lambda y) = z$. Thus a contradiction. Therefore, C is convex. $\$\square$.

A more geometric way of seeing this convexity is to observe that $\{x|f(x) \leq \alpha\}$ is the projection on $\mathbf{R}^n$of the intersection of epi$f$ and the horizontal hyperplane $\{(x,\mu)|\mu = \alpha\}$ in $\mathbf{R}^{n+1}$, so that $\{x|f(x) \leq \alpha\}$ can be regarded as the horizontal cross section of epi$f$.

1.2 The algebra of convex sets

Theorem 3.1: (Sum of conves sets) If C_1 and C_2 are convex sets in $\mathbf{R}^n$, then so is their sum $C_1 + C_2$ where

$$C_1 + C_2 = x_1 + x_2 | x_1 \in C_1, x_2 \in C_2 \tag{-12}$$

Note that here x_1 and x_2 are points in $\mathbf{R}^n$. We would have the vector addition defined on these points. For $(x_{11}, x_{12}, \cdots, x_{1n}) \in C_1$, $(x_{21}, x_{22}, \cdots, x_{2n}) \in C_2$, $(x_{11}, x_{12}, \cdots, x_{1n}) + (x_{21}, x_{22}, \cdots, x_{2n}) = (x_{11}+x_{21}, x_{12}+x_{22}, \cdots, x_{1n}+x_{2n}) \in C_1 + C_2$

1.3 Convex functions

Definition: (Epigraph): Let f be a function whose values are real or $+\infty$ or $-\infty$ and whose domain is a subset S of $\mathbf{R}^n$. The set $\{(x,\mu)|x \in S,\ \mu \in \mathbf{R},\ \mu \geq f(x)\}$ is called the epigraph of f and is denoted as epi$f$.

Definition: (Convex function): We define f to be convex function on S if epi$f$ is convex as a subset of $\mathbf{R}^{n+1}$. That is, we have

$$(1-\lambda)(x,\mu) + \lambda(y,\nu) = ((1-\lambda)x + \lambda y, (1-\lambda)\mu + \lambda\nu) \in \text{epi}f$$

whenever (x,μ) and (y,ν) belong to epi$f$ and $0 \leq \lambda \leq 1$.

Theorem 4.1: Let $f : C \to (-\infty, +\infty]$, where C is a convex set. Then f is convex on C if and only if

$$f((1-\lambda)x+\lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \quad \text{for every } x \text{ and } y \in C,\ 0 \leq \lambda \leq 1$$

Theorem 4.2: Let $f : C \to (-\infty, +\infty]$, where C is a convex set. Then f is convex on C if and only if

$$f((1-\lambda)x+\lambda y) < (1-\lambda)\alpha + \lambda\beta \quad \text{whenever } f(x) < \alpha \text{ and } f(y) < \beta,\ 0 < \lambda < 1$$

1.4 Positively homogenous functions

We will now look at positively homogenous functions. The first question that comes to our mind: Why study about positively homogenous function? what special property do they have that makes them important in the study of convex analysis?

From Wikipedia: In mathematics, a ***homogeneous function*** *is a function with multiplicative scaling behaviour: if the argument is multiplied by a factor, then the result is multiplied by some power of* *this factor. More precisely, if $f : V \to W$ is a function between two vector spaces over a field F, then f is said to be homogeneous of degree $k \in F$ if $f(\alpha x) = \alpha^k f(x)$ for some nonzero $\alpha \in F$ and $x \in V$.*

Theorem 4.4 (Second derivative is non-negative) Let f be a twice continously differentiable real-valued function on an open interval (α, β). Then f is convex if and only if its second derivative is non-negative throughout (α, β).
Proof: 1

1.5 Functional operations

In this section, we look at how to construct new convex functions from functions which are already known to be convex. Often the constructed function is expressed as a constrained infimum, thereby suggesting applications to the theory of extremum problems.

Theorem 5.2: (Sum of convex functions): If f_1 and f_2 are proper convex functions on $\mathbf{R}^n$, then $f_1 + f_2$ is convex.
Proof:

$$\begin{aligned} f_1((1-\lambda)x + \lambda y) &\leq (1-\lambda)f_1(x) + \lambda f_1(y) \\ f_2((1-\lambda)x + \lambda y) &\leq (1-\lambda)f_2(x) + \lambda f_2(y) \quad \text{for every } x \text{ and } y \in C,\ 0 \leq \lambda \leq 1 \end{aligned}$$

Adding the two equations,

$$\begin{aligned} f_1((1-\lambda)x + \lambda y) + f_2((1-\lambda)x + \lambda y) &\leq (1-\lambda)f_1(x) + \lambda f_1(y) + (1-\lambda)f_2(x) + \lambda f_2(y) \\ &= (1-\lambda)(f_1(x) + f_2(x)) + \lambda(f_1(y) + f_2(y)) \end{aligned}$$

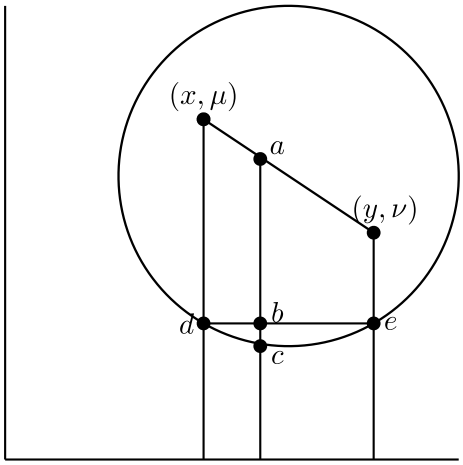
Thus, $f_1 + f_2$ is convex. □

A common device for constructing convex functions on $\mathbf{R}^n$is to construct a convex set F in $\mathbf{R}^{n+1}$and then take the function whose graph is the *lower boundary* of F in the sense of the following theorem.
Theorem 5.3: Let F any convex set in $\mathbf{R}^{n+1}$and let

$$f(x) = \inf\{\mu | (x, \mu) \in F\} \tag{-18}$$

Then f is a convex function on $\mathbf{R}^n$.
Proof: We need to prove that

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$



Given that F is convex, that is ,

$$(1-\lambda)(x,\mu)+\lambda(y,\nu)=((1-\lambda)x+\lambda y,(1-\lambda)\mu+\lambda\nu)\in \text{epi}f$$

whenever (x,μ) and (y,ν) belong to epi$f$ and $0\leq\lambda\leq 1$.

$$f((1-\lambda)x+\lambda y)=\inf\{\zeta|((1-\lambda)x+\lambda y,\zeta)\in F\}$$

Point c in the adjoining figure shows $f((1-\lambda)x+\lambda y)$. That means,

$$f((1-\lambda)x+\lambda y)\leq(1-\lambda)\mu+\lambda\nu$$

$f(x)\leq\mu$ and $f(y)\leq\nu$, therfore we have,

$$f((1-\lambda)x+\lambda y)\leq(1-\lambda)f(x)+\lambda f(y)$$

Thus, f is convex. □

1.5.1 Infimal convolution

Theorem 5.4:(Infimal Convolution) Let f_1, f_2, $\cdots$, f_m be proper convex functions on $\mathbf{R}^n$, and let

$$f(x)=\inf\{f_1(x_1)+\cdots+f_m(x_m)|x_i\in\mathbf{R}^n,x_1+x_2+\cdots+x_m=x\}$$

Then f is a convex function on $\mathbf{R}^n$.

Proof: Let $F_i=\text{epi}f_i$. F_i is convex. From theorem 3.1 we have, $F=F_1+F_2+\cdots+F_m$ is convex. By definition of $f(x)$, we have $(x,\mu)\in F$ if and only if $x_i\in\mathbf{R}^n$ and $\mu_i\in R$ such that $\mu_i\geq f_i(x_i)$, $\mu=\mu_1+\mu_2+\cdots+\mu_m$ and $x=x_1+x_2+\cdots+x_m$. Thus f defined in the theorem is convex from Theorem 5.3. □

The function f in Theorem 5.4 will be denoted by $f_1\Box f_2\Box\cdots\Box f_m$. The operation $\Box$ is called *infimal convolution*. Infimal convolution can be defined as

$$(f\Box g)(x)=\inf\{f(x-y)+g(y)\}$$

It can also be expressed as,

$$(f\Box g)(x)=\inf\{\mu|(x,\mu)\in(\text{epi}f_1+\text{epi}f_2)\}$$

Properties of infimal convolution:

1. Infimal convolution is dual to the operation of addition of convex functions.
2. As an operation on collection of functions, infimal convolution is commutative, associative and convexity preserving. The function $\delta(\cdot|0)$ acts as an identity element for this operation.

1.5.2 Scalar Multiplication

Left scalar multiplication

The operation of non-negative left scalar multiplication preserves convexity, where

$$(\lambda f)(x)=\lambda(f(x))$$

Right scalar multiplication
Right scalar multiplication corresponds to scalar multiplication of epigraphs. For any convex function f on $\mathbf{R}^n$and any λ, $0 \leq \lambda < \infty$, we define $f\lambda$ to be the convex function obtained from Theorem 5.3 with $F = \text{epi} f$. Thus,

$$(f\lambda)(x) = \lambda f(\lambda^{-1}x), \quad \lambda > 0$$

When $\lambda = 0$ we have

$$(f0)(x) = \delta(\cdot|0)$$

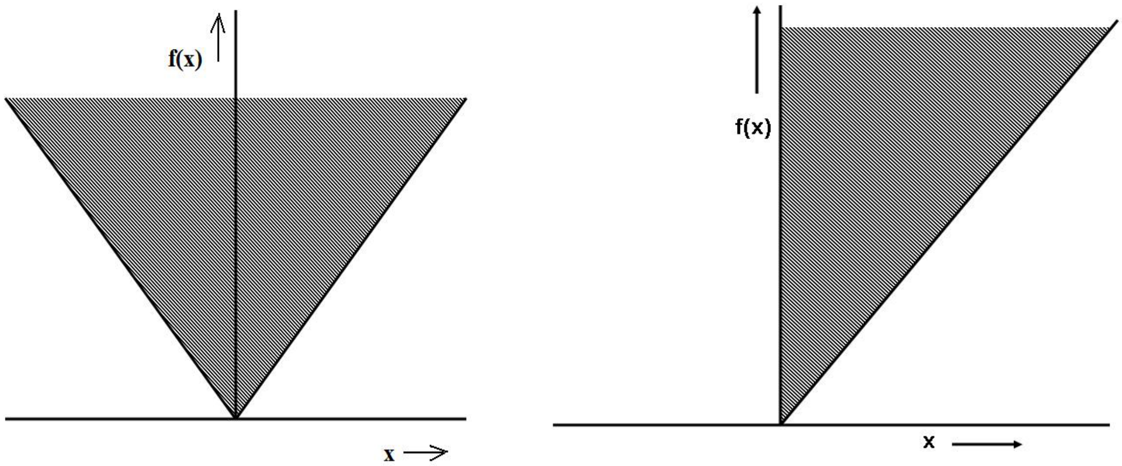
A function f is positively homogeneous if and only if $f\lambda = f$ for every $\lambda > 0$.

Below figure **??** shows that a function of the form

$$\begin{aligned} f(x) &= \alpha x \\ \text{Thus, } (f\lambda)(x) &= \lambda(\alpha(\lambda^{-1}x)) \\ &= \alpha x \\ &= f(x) \end{aligned} \tag{-27}$$

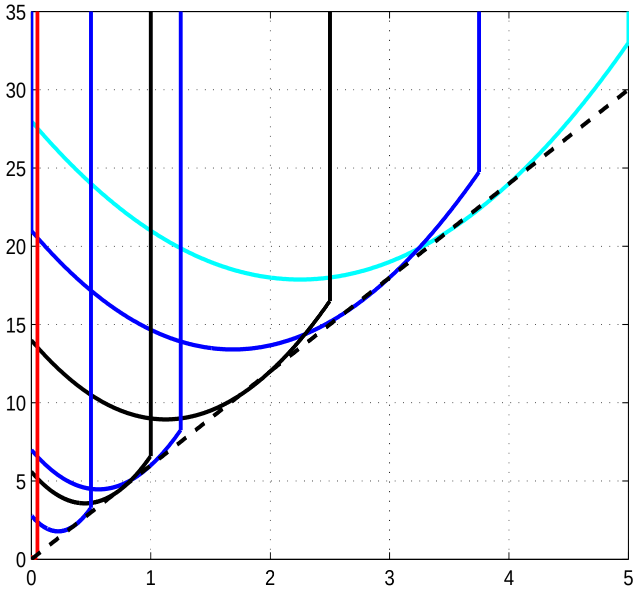
Let h be any convex function in $\mathbf{R}^n$, and let F be the convex cone in $\mathbf{R}^{n+1}$generated by epi$h$. The function obtained by applying Theorem 5.3 to F has as its epigraph a convex cone in $\mathbf{R}^{n+1}$containing the origin. It is the greatest of the positively homogeneous convex functions f such that $f(0) \leq 0$ and $f \leq h$. We call this function the *positively homogeneous convex function generated by* h. Since F consists of the origin and the union of the sets $\lambda(\text{epi}h)$ for $\lambda > 0$, we have

$$f(x) = \inf\{(h\lambda)(x)|\lambda \geq 0\} \tag{-26}$$



(a) (b)

Figure 1: Examples of positively homogenous functions



The adjoining figure illustrates the right multiplication operation. The line in cyan is the function. As we can see, as λ decreases, the function tends to $\delta(\cdot|0)$ as shown in red.

Figure 2: Example illustrating right scalar multiplication

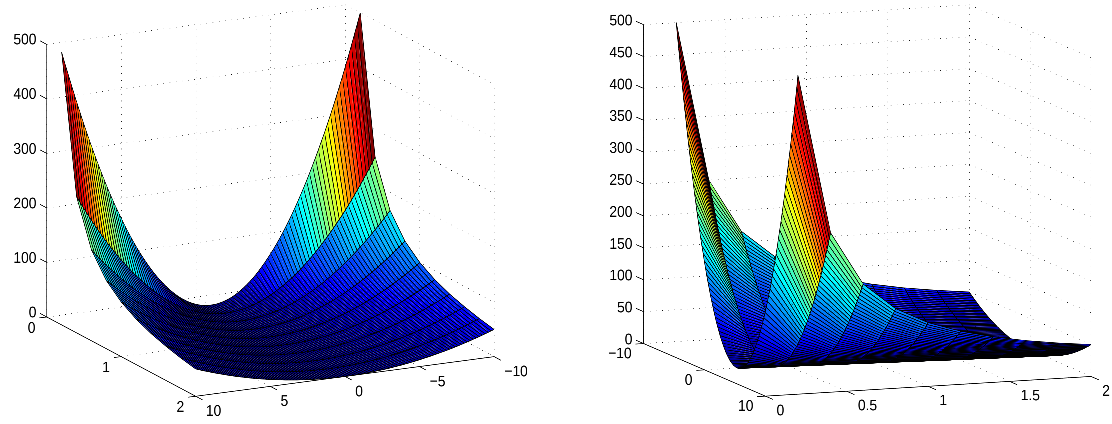
For any proper convex function f on $\mathbf{R}^n$, the function g on $\mathbf{R}^{n+1}$defined by

$$g(\lambda, x) = \begin{cases} (f\lambda)(x) & \text{if } \lambda \geq 0 \\ +\infty & \text{if } \lambda < 0 \end{cases}$$

is positively homogeneous convex function generated by

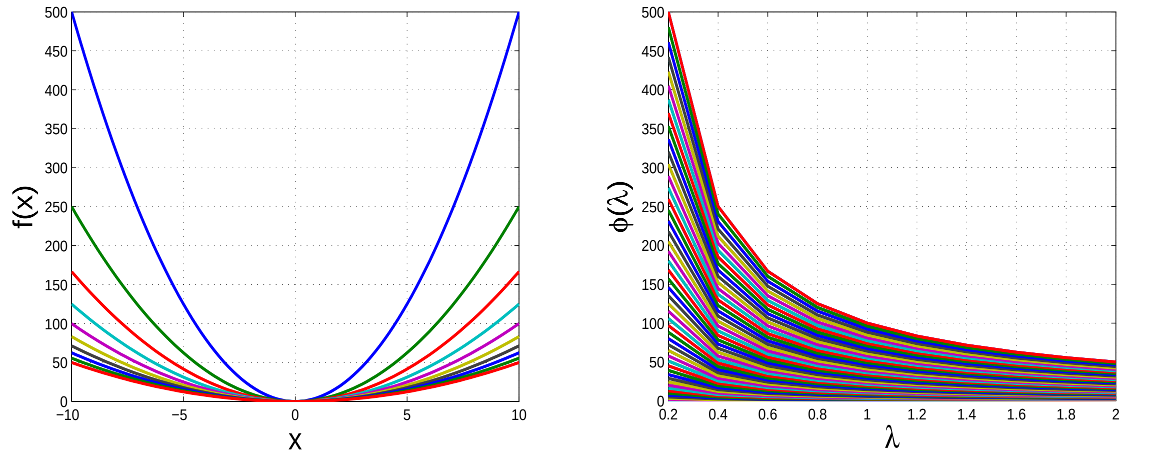
$$h(\lambda, x) = \begin{cases} f(x) & \text{if } \lambda = 1 \\ +\infty & \text{if } \lambda \neq 1 \end{cases}$$

In particular, then, $\phi(\lambda) = (f\lambda)(x)$ is a proper convex function of $\lambda \geq 0$ for any $x \in \text{dom} f$.



(a) (b)

Figure 3: Figure shows the function $g(\lambda, x)$



(a) $(f\lambda)(x)$ as a function of x different values of λ (b) $\phi(\lambda)$ as function of λ for different values of x

Figure 4: Shows the functions of x and λ to be convex

Theorem 5.5 (Pointwise Supremum of functions): The pointwise supremum of an arbitrary collection of convex functions is convex.

Proof: We know that intersection of arbitrary collection of convex sets is convex. Now all we need to prove is that the intersection of the epigraphs of the collection of convex functions is indeed the epigraph of the of the function $f(x) = \sup\{f_i(x)|i \in I\}$. The intersection of epigraphs of the convex functions is the set G given by

$$\begin{aligned} G &= \{(x, \mu)|\mu \geq f_i(x) \quad x \in \mathrm{dom} f_i \text{ for every } i \in I\} \\ &= \{(x, \mu)|\mu \geq \sup\{f_i(x)|x \in \mathrm{dom} f_i, \text{ for every } i \in I\}\} \end{aligned}$$

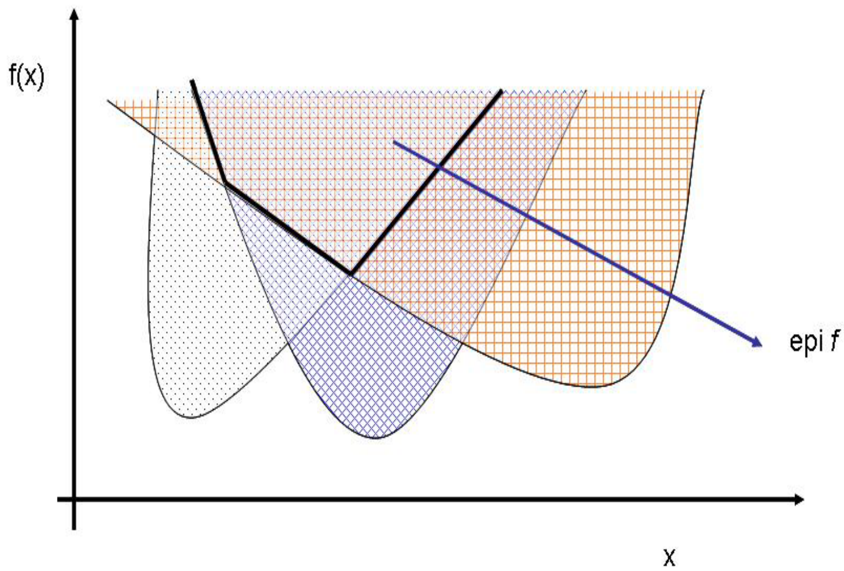


Figure 5: Pointwise supremum of collection of convex functions is convex

Thus, we see that $G = \text{epi}f$ if $f(x) = \sup\{f_i(x)|i \in I\}$. $\square$

2 Topological properties

2.1 Relative interiors of convex sets

We now consider some generic topological properties of convex sets. Let C be a non-empty convex set subset of $\mathbf{R}^n$. The closure of C is also a non-empty convex set. While the interior of C may be empty, it turns out that convexity implies the existence of interior points relative to the affine hull of C.

Theorem : Non-empty interior of convex set

Proof: Given that C is a convex set. Let the dimension of the affine hull of C, $\text{aff}(C)$, be m. This means that there are $m + 1$ points, $\{b_0, b_1, b_2, \cdots, b_m\}$ in C which are affinely independent. This means that we have m vectors, $\{b_1 - b_0, b_2 - b_1, \cdots, b_m - b_0\}$ in C which are linearly independent.

2.2 Recession cones and unboundedness

Discussion: *First question: What is a recession cone? Why study recession cones? What is its relation to optimization problem?*

The notion of recession cone and lineality space can be used to generalize some of the fundamental properties of compact sets to closed convex sets. One such property is that of intersection of nested sequence of nonempty compact sets is non-empty and compact. These properties fail for general closed sets, but it turns out that they hold under certain assumptions involving convexity and directions of recession. These properties translate into important results relating to the existence of solutions of convex optimization problems.

The recession cone of the function can be used to obtain the directions along which function decreases monotonically. It will be shown that a direction that is not a direction of recession of the function is a direction of eventual continuous ascent of the function.

---

3 Separation Theorems

4 Conjugates of convex function

There are two ways of viewing a classical curve or surface like a conic, either as a locus of points or as an envelope of tangents. This fundamental duality enters the theory of convexity in a slightly different form: a closed convex set in $\mathbf{R}^n$is the intersection of the closed half-spaces which contain it.

The definition of the conjugate of a function grows naturally out of the fact that the epigraph of a closed proper convex function on $\mathbf{R}^n$is the intersection of the closed half-spaces in $\mathbf{R}^{n+1}$which contain it.
The hyperplanes in $\mathbf{R}^{n+1}$can be represented by means of the linear functions (or affine functions : affine functions are functions which are finite, convex and concave.) on $\mathbf{R}^{n+1}$, and these can in turn be represented in the form

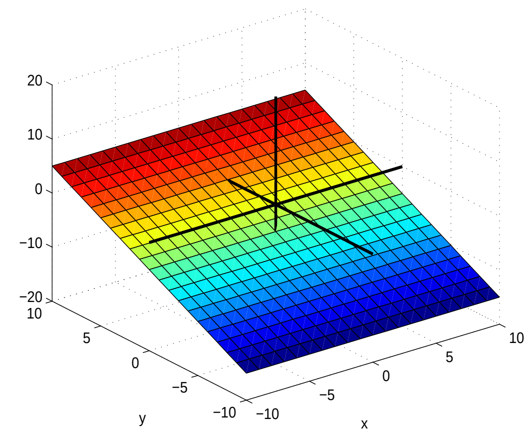
$$(x, \mu) \to \langle x, b \rangle + \mu \beta_0$$

where $b \in \mathbf{R}^n$, $\beta_0 \in \mathbf{R}$. The above equation means that the hyperplane would be of the form

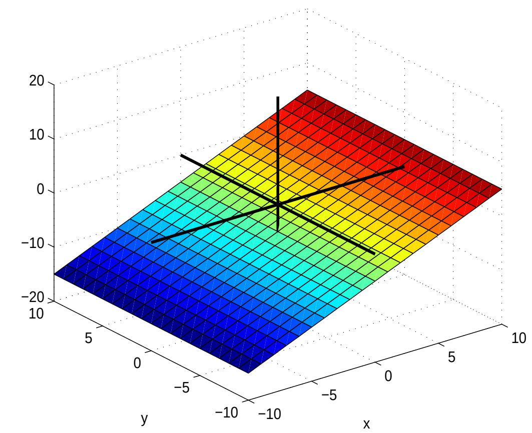
$$\begin{aligned} g(x, \mu) &= \beta \\ &= \langle x, b \rangle + \mu \beta_0 = \beta \end{aligned}$$

Observe that $x \in \mathbf{R}^n$. We infact are taking about a convex function $f : \mathbf{R}^n \to (-\infty, +\infty]$, $\mu = f(x)$. We are looking at the epigraph of the convex function which is a set consisting of points in $\mathbf{R}^{n+1}$. We are trying to find hyperplanes and half-spaces such that the resultant of the intersection of these half-spaces is the epigraph of this function f.

Discussion: *Consider the following example. Consider the function $f : \mathbf{R}^2 \to \mathbf{R}$. Now the above hyperplane is nothing but a hyperplane with the normal $n = [b_1\, b_2\, -1]$. $x = [x_1\, x_2]$. Consider $b = [b_1\, b_2] = [0\,1]$. Let the co-ordinate system have it basis along $\hat{x_1}$, $\hat{x_2}$ and $\hat{\mu}$. When $b = [b_1\, b_2] = [0\,1]$, the resultant vector n will be along the $\hat{x_2}$-$\hat{\mu}$ plane. The hyperplane will be as shown below. With different values of β the hyperplane shifts parallelly as shown. Then we have*



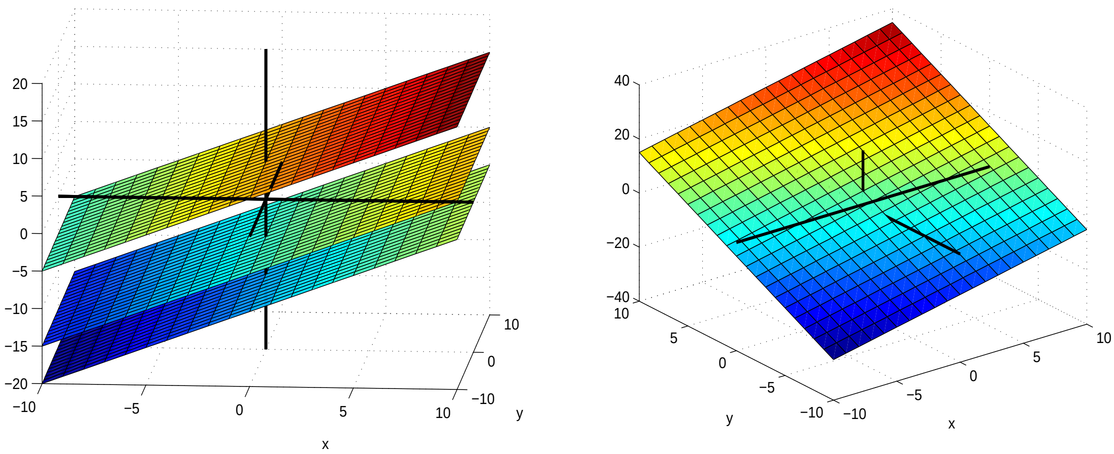
(a) $b = [b_1\, b_2] = [0\,1]$



(b) $b = [b_1\, b_2] = [1\,0]$

Figure 6: Hyperplane

$b_1 \neq 0$ and $b_2 \neq 0$, we get a plane as shown below.



(a) $b = [b_1\, b_2] = [1\, 0]$, different β's

(b) $b = [b_1\, b_2] = [1\, 2]$

Figure 7: Hyperplane

Theorem 12.1: A closed convex function f is the pointwise supremum of the collection of all affine functions h such that $h \leq f$.