Multi-User Information Theory Jan 17, 2012

Lecture 10

Lecturer:Dr. Haim Permuter *Scribe: Wasim Huleihel*

I. Conjugate Function

In the previous lectures, we discussed about convex set, convex functions (properties and examples), and we present some operations that preserve convexity. This lecture we will continue the discussion about convex functions; definition and examples of conjugate functions will be given, and we will present some general convex optimization problems.

Definition 1 (Conjugate function) Let $f : \mathbb{R}^n \to \mathbb{R}$. The function $f^* : \mathbb{R}^n \to \mathbb{R}$ defined as

$$f^*(y) \triangleq \sup_{x \in \text{dom} f} \left(y^T x - f(x)\right), \tag{1}$$

is called the *conjugate* function of f.

This definition is illustrated in Fig. 1.

Remark 1 The domain of the conjugate function is given by

$$\text{dom} f^* = \left\{ y \in \mathbb{R}^n : \sup_{x \in \text{dom} f} \left(y^T x - f(x)\right) < \infty \right\}. \tag{2}$$

Remark 2 The conjugate function, f^*, is a convex function. This can be easily verified using that fact that the supremum of a set of convex functions (in our case, for a fixed x, the difference $y^T x - f(x)$ is a linear function of y and hence the difference is convex) is convex function. Note that this is true whether or not f is convex.

Remark 3 $y^T x$ stands for inner product in finite space dimensions. In infinite dimensional vector spaces, the conjugate function is defined via an appropriate inner product.

This operation play an important role in many applications such as duality. In the following, we present some examples of conjugate functions.

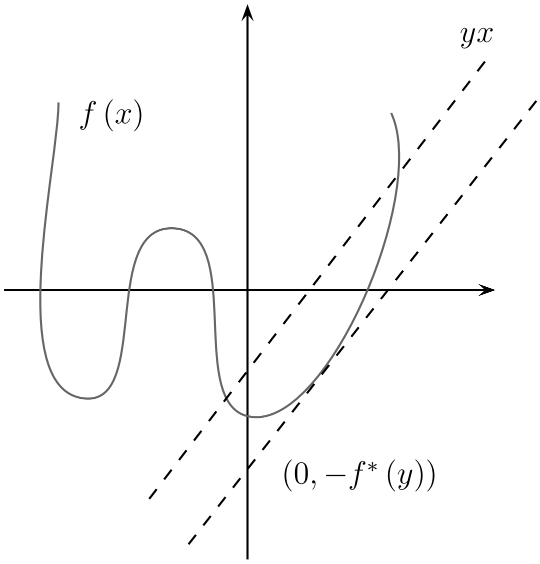

Fig. 1. The conjugate function $f^*(y)$ is the maximum gap between the linear function yx and the function $f(x)$.

Example 1 (Affine function) $f(x) = ax + b$. By definition, the conjugate function is given by $f^*(y) = \sup_x (yx - ax - b)$. As a function of x, the difference is bounded iff $y - a = 0$. The conjugate function is given $f^*(y) = -b$, and the domain is $\mathrm{dom} f^* = \{a\}$.

Example 2 (Negative logarithm) $f(x) = -\log x$, with $\mathrm{dom} f = \mathbb{R}_{++}$. By definition, the conjugate function is given by $f^*(y) = \sup_x (yx + \log x)$. As a function of x, the difference is bounded iff $y < 0$, and reaches its maximum at $x = -\frac{1}{y}$. The conjugate function is given $f^*(y) = -\log(-y) - 1$, and the domain is $\mathrm{dom} f^* = -\mathbb{R}_{++}$.

Example 3 (Exponential) $f(x) = e^x$. By definition, the conjugate function is given by $f^*(y) = \sup_x (yx - e^x)$. As a function of x, the difference is bounded iff $y > 0$, and reaches its maximum at $x = \ln y$. The conjugate function is given $f^*(y) = y \ln y - y$, and the domain is $\mathrm{dom} f^* = \mathbb{R}_{++}$.

Example 4 (Entropy) $f(x) = x \log x$, with $\mathrm{dom} f = \mathbb{R}_{++}$. By definition, the conjugate function is given by $f^*(y) = \sup_x (yx - x \log x)$. As a function of x, the difference is bounded for all y, and reaches its maximum at $x = e^{(y-1)}$. The conjugate function is given $f^*(y) = e^{(y-1)}$, and the domain is $\mathrm{dom} f^* = \mathbb{R}$. So the conjugate of entropy is related to the exponential family.

Example 5 (Log-sum-exp) $f(x) = \log \sum_{i=1}^{n} e^{x_i}$, with dom$f = \mathbb{R}^n$. By definition, the conjugate function is given by $f^*(y) = \sup_x \left(yx - \log \sum_{i=1}^{n} e^{x_i}\right)$. By setting the gradient with respect to x to zero, the critical points are given by

$$y_i = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}} \; i = 1, \ldots, n. \tag{3}$$

The domain of f^* is dom$f^* = \{y \in \mathbb{R}^n : 0 \le y_i \le 1, \; \sum_{i=1}^{n} y_i = 1\}$. The conjugate function is given $f^*(y) = \sum_{i=1}^{n} y_i \log y_i$.

Next, we present two interesting properties of conjugate function.

Properties of conjugate function

- $f^*(y) + f(x) \ge x^T y$.
- Conjugate of the conjugate: if f is convex function, and its epigraph is closed set, then $f^{**} = f$.

II. Convex Optimization Problems

The general optimization problem is given by

$$\begin{aligned} &\min_x f_0(x) \\ &\text{s.t. } f_i(x) \le 0; \quad i = 1, \ldots, m \\ &\qquad h_j(x) = 0; \quad j = 1, \ldots, p, \end{aligned} \tag{4}$$

where $f_0(x)$,$\{f_i(x)\}_{i=1}^{m}$ are convex functions, and $\{h_j(x)\}_{j=1}^{p}$ are affine functions. The optimization problem (4) describe the problem of finding an x that minimizes the objective f_0 subject to (s.t.) the constraints $f_i(x) \le 0; \; i = 1, \ldots, m$, and $h_j(x) = 0; \; j = 1, \ldots, p$. In the following, we present some classes of optimization problems that can be efficiently solved using Matlab.

- *Linear Programming*: linear programming is a technique for the optimization of a linear objective function, subject to linear equality and linear inequality constraints.

Its feasible region (the space of all possible solutions), is a convex polyhedron. Mathematically, the optimization problem in this case can be in the following form

$$\begin{aligned} &\min_x c^T x + d \\ &\text{s.t. } Gx \leq h \\ &\quad\;\; Ax = b. \end{aligned} \tag{5}$$

- *Linear Fraction*: linear fraction optimization problem can be written as

$$\begin{aligned} &\min_x \frac{c^T x + d}{e^T x + f} \\ &\text{s.t. } Gx \leq h \\ &\quad\;\; Ax = b \\ &\quad\;\; e^T x + f > 0. \end{aligned} \tag{6}$$

Let us transform this optimization problem to a linear one (5). This can be easily done by introducing auxiliary variables defined as

$$z = \frac{1}{e^T x + f},$$

and

$$y = \frac{x}{e^T x + f}.$$

Using these new variables the objective function become $c^T y - dz$, and the constraints become

$$\begin{aligned} &Gx \leq h \Rightarrow Gy \leq hz \\ &Ax = b \Rightarrow Ay = bz \\ &e^T x + f > 0 \Rightarrow z \geq 0 \\ &e^T y + fz = 1 \Rightarrow \text{New constraint.} \end{aligned}$$

Therefore the optimization problem is given by

$$\min_{y,z} c^T y - dz$$

$$\text{s.t. } Gy \leq hz$$
$$Ay = bz$$
$$z \geq 0,\ \ e^T y + fz = 1. \tag{7}$$

- *Quadratic Optimization Problem*: the quadratic optimization problem is given by

$$\min_x \frac{1}{2} x^T P x + q^T x + r$$
$$\text{s.t. } Gx \leq h$$
$$Ax = b. \tag{8}$$

  where $P \in \mathbb{S}^n_+$.

- *Second Order Cone Programming*: the second order cone optimization problem is given by

$$\min_x f^T x$$
$$\text{s.t. } \ \|A_i x + b\|_2 \leq c_i^T x + b_i; \quad i = 1, \ldots, n$$
$$Fx = g. \tag{9}$$

- *Geometric Programming*: this class of problems is not convex, but we will transform it to a convex one.

  Definition 2 The function $f(x) = cx_1^{q_1} \cdot x_2^{q_2} \cdot \ldots \cdot x_n^{q_n}$, where $c \geq 0$, is called *monomial*.

  Definition 3 The function $f(x) = \sum_{k=1}^K c_k x_1^{q_{1,k}} \cdot x_2^{q_{2,k}} \cdot \ldots \cdot x_n^{q_{n,k}}$, is called *posynomial*.

  The Geometric optimization problem is given by

$$\min_x f_0(x)$$
$$\text{s.t. } f_i(x) \leq 0; \quad i = 1, \ldots, m$$
$$h_j(x) = 0; \quad j = 1, \ldots, p, \tag{10}$$

  where $f_0(x)$,$\{f_i(x)\}_{i=1}^m$ are posynomial functions, and $\{h_j(x)\}_{j=1}^p$ are monomial functions.

Example 6 Consider the following optimization problem

$$
\begin{aligned}
&\min_{x,y,z} \frac{x}{y} \\
&\text{s.t. } 2 \leq x \leq 3 \\
&\quad x^2 + \frac{3y}{z} \leq \sqrt{y} \\
&\quad \frac{x}{y} = z^2.
\end{aligned}
$$

is a Geometric optimization problem.

Geometric programs are not (in general) convex optimization problems, but they can be transformed to convex problems by a change of variables and a transformation of the objective and constraint functions. We will use the variables defined as $x_i = e^{y_i}$. Using this transformation, the Geometric optimization problem given in (10), can be rewritten as

$$
\begin{aligned}
&\min_{y} \log \left[\sum_{k=1}^{K_0} \exp\left(\sum_m a_{0,k,m} y_m + b_{0,k} \right) \right] \\
&\text{s.t. } \log \left[\sum_{k=1}^{K_i} \exp\left(\sum_m a_{i,k,m} y_m + b_{i,k} \right) \right] \leq 0; \quad i = 1, \ldots, m \\
&\quad \sum_m c_{j,m} y_m + d_j = 0; \quad j = 1, \ldots, p.
\end{aligned} \tag{11}
$$

Since the objective function, and the first constraint are convex (log-sum exponential function), and the second constraint is affine, this problem is convex optimization one.

- *Semi-Definite Programming (SDP)* the optimization problem is given by

$$
\begin{aligned}
&\min_{x} e^T x \\
&\text{s.t. } x_1 F_1 + \ldots + x_n F_n \preceq 0 \\
&\quad Ax = b,
\end{aligned} \tag{12}
$$

References

[1] S. Boyd, *Convex Optimization.* Cambridge University Press , 2004.