Mathematical methods in communication — January 10th, 2012

Lecture 9

Lecturer: Haim Permuter — *Scribe: Moti Teitel*

I. INRODUCTION

Lemma 1 f is convex iff $g(t) = f(x + tv)$ is convex in t for all v.

Exercise: Prove Lemma 1.

For example, we examine $\log\det \mathbf{X}$, where $\mathbf{X} \in S_{++}^{N}$. Assume that $\mathbf{V} \in S^n$. Then:

$$\begin{aligned}
g(t) &= \log|\mathbf{X} + t\mathbf{V}| \overset{(a)}{=} \log\left|\mathbf{X}^{1/2}\mathbf{X}^{1/2} + t\mathbf{V}\right| \\
&= \log\left|\mathbf{X}^{1/2}\left(\mathbf{I} + t\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2}\right)\mathbf{X}^{1/2}\right| \\
&= \log\left|\mathbf{X}^{1/2}\left(\mathbf{I} + t\mathbf{U}\Lambda\mathbf{U}^T\right)\mathbf{X}^{1/2}\right| \\
&= \log\left|\mathbf{X}^{1/2}\mathbf{U}\left(\mathbf{I} + t\Lambda\right)\mathbf{U}^T\mathbf{X}^{1/2}\right| \\
&= \log\left|\mathbf{X}^{1/2}\right||\mathbf{U}||\mathbf{I} + t\Lambda|\left|\mathbf{U}^T\right|\left|\mathbf{X}^{1/2}\right| \\
&\overset{(b)}{=} \log\left|\mathbf{X}^{1/2}\right||\mathbf{I} + t\Lambda|\left|\mathbf{X}^{1/2}\right| \\
&= \log\left|\mathbf{X}^{1/2}\right| + \log|\mathbf{I} + t\Lambda| + \log\left|\mathbf{X}^{1/2}\right| \\
&= \log|\mathbf{X}| + \log|\mathbf{I} + t\Lambda| \\
&\overset{(c)}{=} \log|\mathbf{X}| + \sum \log(1 + t\lambda_i)
\end{aligned} \tag{1}$$

which is concave in t, Therefore $\log\det \mathbf{X}$ is concave in $\mathbf{X}$.

The equality marked by (a) is due to the fact that $\mathbf{X} \in S_{++}^{n}$, so we can write $\mathbf{X} = \mathbf{X}^{1/2}\mathbf{X}^{1/2}$.

Equality (b) holds since $\mathbf{X}, \mathbf{V} \in S^n$ so $\mathbf{X}^{-1/2}\mathbf{V}\mathbf{X}^{-1/2} \in S^n$ as well, and is therefore diagonalizable.

Equality (c) follows from the fact that $\mathbf{U}$ is unitary, so $\det \mathbf{U} = 1$.

Equality (d) follows from the fact that Λ is diagonal.

II. Operations that preserve convexity

1) **Non-negative weighted sum**: if $\{f_i\}_{i=1}^{N}$ are convex, and $\{w_i\}_{i=1}^{N}$ are non-negative scalars, then $\sum_i w_i f_i$ is convex. This is also true for countably infinite suns, and for integrals: If for all $y \in A$ we have $f(x, y)$ convex in x, and $w(y) \geq 0$ for all $y \in A$, then

$$\int_A f(x, y)\, w(y)\, dy \tag{2}$$

is convex in x.

2) **pointwise maximum**: if f_1, f_2 are convex, then $\max(f_1, f_2)$ is also convex.

Proof:

$$\begin{aligned} f\left(\theta x + \bar{\theta} y\right) &= \max\left(f_1\left(\theta x + \bar{\theta} y\right), f_2\left(\theta x + \bar{\theta} y\right)\right) \\ &\leq \max\left(\theta f_1(x) + \bar{\theta} f_1(y), \theta f_2(x) + \bar{\theta} f_2(y)\right) \\ &\leq \max\left(\theta f_1(x), \theta f_2(x)\right) + \max\left(\bar{\theta} f_1(y), \bar{\theta} f_2(y)\right) \\ &= \theta f(x) + \bar{\theta} f(y) \end{aligned} \tag{3}$$

■

Using this property, we may conclude that the intersection of convex sets is convex. In order to show this, recall that a function is convex iff it's epigraph is a convex set. Let f_1 and f_2 be the two functions whose epigraphs are the convex sets A and B respectivally. Since the intersection of the two convex sets A, B is the epigraph of $\max(f_1, f_2)$, which is convex, it is a convex set.

3) **composition with an affine mapping**: suppose $\mathbf{P} : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{n+m}, \mathbf{b} \in \mathbb{R}^n$. Define $g : \mathbb{R}^m \to \mathbb{R}$ by $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$. The domain of g is $\mathrm{dom} g = \{\mathbf{x} : \mathbf{Ax} + \mathbf{b} \in \mathrm{dom} f\}$. If f is convex then g is convex for all $\mathbf{A}, \mathbf{b}$.

Proof:

$$\begin{aligned} g\left(\theta \mathbf{x} + \bar{\theta} \mathbf{y}\right) &= f\left(\mathbf{A}\left(\theta \mathbf{x} + \bar{\theta} \mathbf{y}\right) + \mathbf{b}\right) = f\left(\theta \mathbf{Ax} + \bar{\theta} \mathbf{Ay} + \mathbf{b}\right) \\ &= f\left(\theta(\mathbf{Ax} + \mathbf{b}) + \bar{\theta}(\mathbf{Ay} + \mathbf{b})\right) \leq \theta f(\mathbf{Ax} + \mathbf{b}) + \bar{\theta} f(\mathbf{Ay} + \mathbf{b}) \\ &= \theta g(\mathbf{x}) + \bar{\theta} g(\mathbf{y}) \end{aligned} \tag{4}$$

■

4) **Scalar composition** (can be extended to a vector space): Given $g : \mathbb{R} \rightarrow \mathbb{R}, h : \mathbb{R} \rightarrow \mathbb{R}$, define $f = h\left(g\left(x\right)\right)$.
   a) If h is a convex function, non-decreasing, and g is convex, then f is convex.
   b) If h is a convex function, non-increasing, and g is concave, then f is convex.
   c) If h is a concave function, non-increasing, and g is convex, then f is concave.
   d) If h is a concave function, non-decreasing, and g is concave, then f is concave.

   For example, if g is convex, $e^{g(x)}$ is convex. If g is concave and positive, $\log g\left(x\right)$ is concave.

   Proof: We will prove (a): Since h is non-decreasing, we have $h' \geq 0$. Since h is convex, we have $h'' \geq 0$. Since g is convex, we have $g'' \geq 0$. So,

$$
\begin{aligned}
\left[h\left(g\right)\right]' &= h'\left(g\right) g' \qquad (5)\\
\left[h\left(g\right)\right]'' &= \left[h'\left(g\right) g'\right]' = h''\left(g\right)\left(g'\right)^2 + h'\left(g\right) g'' \geq 0
\end{aligned}
$$

■

   The rest of the properties are proven similarly.

5) **Vector composition**: $f\left(x\right) = h\left(g\left(x\right)\right) = h\left(g_1\left(x\right), g_2\left(x\right), \ldots, g_k\left(x\right)\right)$, where $h : \mathbb{R}^k \rightarrow \mathbb{R}, g : \mathbb{R}^n \rightarrow \mathbb{R}^k, g_i : \mathbb{R}^n \rightarrow \mathbb{R} \;\; i = 1, 2, \ldots, k$.
   a) If h is a convex function, non-decreasing, and g_i are convex for all i, then f is convex.
   b) If h is a convex function, non-increasing, and g_i are concave for all i, then f is convex.
   c) If h is a concave function, non-increasing, and g_i are convex for all i, then f is concave.
   d) If h is a concave function, non-decreasing, and g_i are concave for all i, then f is concave.

   If h is convex, non-decreasing in each element, and g_i are concave for all i, then f is convex.

For $n = 1$, for example, we have

$$f'(x) = \nabla h^T(g(x))\, g'(x), \quad g'(x) = (g_1'(x), g_2'(x), \ldots, g_k'(x)) \tag{6}$$
$$f''(x) = (g'(x))^T \nabla^2 h(g(x))\, g'(x) + \nabla h^T(g(x))\, g''(x)$$

So if the conditions hold we have that $f'' \geq 0$.

The other combinations follow similarly.

6) If $f(x, y)$ is a convex function in $(x, y) \in C$, where C is a convex set, then

$$g(x) = \inf_{y:(x,y)\in C} f(x, y) \tag{7}$$

is a convex function.

Proof: From the definition of the infimum we have

$$\forall \epsilon > 0,\ \forall x_1, x_2\ \exists y_1, y_2: \quad f(x_i, y_i) \leq g(x_i) + \epsilon\ i = 1, 2. \tag{8}$$

We may then write, for all $\epsilon > 0$,

$$g\left(\theta x_1 + \bar{\theta} x_2\right) = \inf_y f\left(\theta x_1 + \bar{\theta} x_2, y\right) \overset{(a)}{\leq} f\left(\theta x_1 + \bar{\theta} x_2, \theta y_1 + \bar{\theta} y_2\right) \tag{9}$$
$$\leq \theta f(x_1, y_1) + \bar{\theta} f(x_2, y_2) \overset{(b)}{\leq} \theta g(x_1) + \bar{\theta} g(x_2) + 2\epsilon$$

■

where (a) holds as $\inf_y$ must achieve a lower (or equal) value that any specific y, in particular than for $\theta y_1 + \bar{\theta} y_2$, and (b) follows from Equation 8.

7) **Perspective transformations**: If $f(x)$ is a convex function, then $g(x, t) = tf\left(\frac{f}{t}\right)$ is also convex in (x, t) for $t \geq 0$.

Proof:

$$g\left(\theta(x_1, t_1) + \bar{\theta}(x_2, t_2)\right) = \left(\theta t_1 + \bar{\theta} t_2\right) f\left(\frac{\theta x_1 + \bar{\theta} x_2}{\theta t_1 + \bar{\theta} t_2}\right) \tag{10}$$
$$= \left(\theta t_1 + \bar{\theta} t_2\right) f\left(\frac{\theta t_1 \frac{x_1}{t_1} + \bar{\theta} t_2 \frac{x_2}{t_2}}{\theta t_1 + \bar{\theta} t_2}\right)$$
$$\leq \left(\theta t_1 + \bar{\theta} t_2\right)\left[\frac{\theta t_1}{\theta t_1 + \bar{\theta} t_2} f\left(\frac{x_1}{t_1}\right) + \frac{\bar{\theta} t_2}{\theta t_1 + \bar{\theta} t_2} f\left(\frac{x_2}{t_2}\right)\right]$$
$$= \theta t_1 f\left(\frac{x_1}{t_1}\right) + \bar{\theta} t_2 f\left(\frac{x_2}{t_2}\right) = \theta g(x_1, t_1) + \bar{\theta} g(x_2, t_2)$$

■

For example, define $P = [p_1, p_2, \ldots, p_n]$ and $Q = [q_1, q_2, \ldots, q_n]$ to be probability distributions, and examine $D(P||Q) = \sum_i p_i \log \frac{p_i}{q_i} = -\sum p_i \log \frac{q_i}{p_i}$.
Since $f(x) = -\log x$ is convex, we have $g(x,t) = tf\left(\frac{f}{t}\right) = -t \log \frac{x}{t}$ convex as well. Since the sum of convex functions is convex, we may define $D(X,T) = \sum_i g(x_i, t_i) = -\sum_i t_i \log \frac{x_i}{t_i}$ convex. By substituting $X = Q, T = P$, we have that $D(P||Q) = -\sum p_i \log \frac{q_i}{p_i}$ is convex.