Homework 3

Submit answers for problems 1–6. Problems 7–10 are included as practice problems.

1. *Generalized mean inequalities.* Suppose $w \in \mathbf{R}^n_{++}$ is a positive weight vector, normalized to satisfy $\mathbf{1}^T w = 1$. For positive $x \in \mathbf{R}^n_{++}$, define the generalized mean $f_r(x)$, parametrized by $r \in \mathbf{R}$, as

$$f_0(x) = \prod_{k=1}^n x_k^{w_k}, \qquad f_r(x) = (\sum_{k=1}^n w_k x_k^r)^{1/r} \quad \text{if } r \neq 0.$$

For $r = 0$ this is the weighted geometric mean. Examples for $r \neq 0$ are the weighted arithmetic mean ($r = 1$) and the weighted harmonic mean ($r = -1$).

We show that $f_r(x) \leq f_s(x)$ if $r < s$. This inequality includes as special cases the arithmetic–geometric mean inequality ($r = 0$, $s = 1$) and the arithmetic–harmonic mean inequality ($r = -1$, $s = 1$).

(a) First consider $r = 0 < s$. If we take the logarithm of both sides of the inequality $f_0(x) \leq f_s(x)$, we get the equivalent inequality

$$\sum_{k=1}^n w_k \log x_k \leq \frac{1}{s} \log \sum_{k=1}^n w_k x_k^s.$$

Show that this follows from concavity of the logarithm.

(b) Next consider $0 < r < s$. Show that the inequality $f_r(x) \leq f_s(x)$ follows from convexity of the function $t^{s/r}$ for $t > 0$.

The remaining cases ($r < s = 0$ and $r < s < 0$) follow from (a), (b) and the fact that $f_r(x) = 1/f_{-r}(y)$ where $y_i = 1/x_i$ for $i = 1, \ldots, n$.

(c) What are the conditions on x to have equality $f_r(x) = f_s(x)$ when $r \neq s$?

2. Suppose the functions $g : \mathbf{R} \to \mathbf{R}$ and $f : \mathbf{R} \to \mathbf{R}$ are convex. Show that the function $h : \mathbf{R}^n \to \mathbf{R}$ defined as

$$h(x) = \inf_z \left(g(z) + \sum_{i=1}^n p_i f(x_i - z) \right)$$

is convex, where $p_i > 0$ for $i = 1, \ldots, n$. The variable z in the minimization is a scalar.

Remark. The function h arises in stochastic optimization. Suppose X is a discrete random variable which takes the value x_i with probability p_i. The random variable X is the quantity of a product you will purchase in the future. The price for purchasing the quantity X is $f(X)$. You have the option of preordering an amount z at a price $g(z)$. As a rational consumer, you choose z by minimizing

$$g(z) + \mathbf{E}\, f(X - z) = g(z) + \sum_{i=1}^{n} p_i f(x_i - z).$$

3. Exercise A3.5 (a,b).

4. Exercise A3.48 (a,b). The function in part (b) can be treated as an application of the perspective composition rule of A3.5(b). Another short proof is to show that the epigraph of f is a convex set.

5. Exercise A3.44.

6. Exercise A4.17.

7. Exercise A3.21.

   Solution.

   (a) Using the result mentioned in the problem, we can express a doubly stochastic matrix S as a convex combination $S = \sum_k \theta_k P_k$ of permutation matrices P_k. From convexity and symmetry of f,

   $$f(Sx) = f(\sum_k \theta_k P_k x) \leq \sum_k \theta_k f(P_k x) = \sum_k \theta_k f(x) = f(x).$$

   The inequality is Jensen's inequality (convexity of f). The second equality follows from the symmetry of f.

   (b) The diagonal elements of $Y = Q\,\mathbf{diag}(\lambda)Q^T$ are given by

   $$Y_{ii} = \sum_{j=1}^{n} Q_{ij}^2 \lambda_j.$$

   A matrix Q is orthogonal if $QQ^T = Q^TQ = I$. From $QQ^T = I$, we see that $\sum_{j=1}^{n} Q_{ij}^2 = 1$ for all i. From $Q^TQ = I$, we have $\sum_{i=1}^{n} Q_{ij}^2 = 1$ for all j. Therefore the matrix with elements $S_{ij} = Q_{ij}^2$ is doubly stochastic.

   (c) Combining the results in parts (a) and (b), we conclude that for any symmetric X, the inequality

   $$f(\mathbf{diag}(X)) \leq f(\lambda(X))$$

   holds. Also, if V is orthogonal, $\lambda(X) = \lambda(V^TXV)$. Therefore

   $$f(\mathbf{diag}(V^TXV)) \leq f(\lambda(X))$$

for all orthogonal V. Moreover, this inequality holds with equality if $V = Q$, where Q is the matrix of eigenvectors of X. Hence

$$f(\lambda(X)) = \sup_{V \in \mathcal{V}} f(\mathbf{diag}(V^T X V)).$$

For fixed V, the function $f(\mathbf{diag}(V^T X V))$ is the composition of f with a linear function of X, and therefore convex. We have shown that $f(\lambda(X))$ is the supremum of a family of convex functions of X. Therefore $f(\lambda(X))$ is a convex function of X.

8. Exercise A3.31.

   Solution.

   (a) Follows from the composition rules.

   (b)

   $$\begin{aligned} f^*(y) = \sup_x (y^T x - h(\|x\|_2) &= \sup_{t \geq 0} \sup_{\|x\|_2 = t} (y^T x - h(t)) \\ &= \sup_{t \geq 0} (t\|y\|_2 - h(t)) \\ &= \sup_t (t\|y\|_2 - h(t)) \\ &= h^*(\|y\|_2). \end{aligned}$$

   The third line follows because $h(t) = h(0)$ for $t \leq 0$.

   (c) The conjugate of $h(t)$ is

   $$h^*(s) = \sup_t (st - h(t)) = \begin{cases} (1/q)s^q & s \geq 0 \\ +\infty & s < 0, \end{cases}$$

   where $q = p/(p-1)$. To see this, we distinguish the two cases. If $s < 0$, the function $st - h(t)$ goes to infinity for $t \to -\infty$. If $s \geq 0$, it reaches a maximum where

   $$s - h'(t) = s - t^{p-1} = 0.$$

   Substituting the optimal $t = s^{1/(p-1)}$ in $st - h(t)$ gives $(1/q)s^q$.

9. Exercise T3.23 (a).

   Solution. This is the perspective function of the convex function $\|x\|_p^p = |x_1|^p + \cdots + |x_n|^p$.

10. Exercise A3.20 (c).

    Solution. By making a change of variables $t = 1/\alpha$, we can write this as

    $$f(x) = \inf_{t > 0} (t(g(y + x/t) - g(y))).$$

    This is the infimum over t of the perspective of the convex function

    $$h(x) = g(y + x) - g(y)).$$

    The perspective is convex, jointly in x, t. Partial minimization over t results in a convex function of x.