Hindawi Publishing Corporation
Journal of Inequalities and Applications
Volume 2008, Article ID 717614, 14 pages
doi:10.1155/2008/717614

Research Article

A Refinement of Jensen's Inequality for a Class of Increasing and Concave Functions

Ye Xia

Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL 32611-6120, USA

Correspondence should be addressed to Ye Xia, yx1@cise.ufl.edu

Received 23 January 2008; Accepted 9 May 2008

Recommended by Ondrej Dosly

Suppose that $f(x)$ is strictly increasing, strictly concave, and twice continuously differentiable on a nonempty interval I--, and f-(x) is strictly convex on I. Suppose that x_k-$[a,b]$-I, where $0 < a < b$, and p_k-0 for $k = 1$, -, n, and suppose that -$_{k=1}^{n} p_k = 1$. Let $\bar{x}$= -$_{k=1}^{n} p_k x_k$, and -^{2} = -$_{k=1}^{n} p_k (x_k$- $\bar{x})^2$. We show -$_{k=1}^{n} p_k f(x_k)$-$f(\bar{x}$ --$_1$-$^2)$, -$_{k=1}^{n} p_k f(x_k)$-$f(\bar{x}$ --$_2$-$^2)$, for suitably chosen -$_1$ and -$_2$. These results can be viewed as a refinement of the Jensen's inequality for the class of functions specified above. Or they can be viewed as a generalization of a refined arithmetic mean-geometric mean inequality introduced by Cartwright and Field in 1978. The strength of the above result is in bringing the variations of the x_k's into consideration, through -^{2}.

1. Main theorem

The goal is to generalize the following refinement of the arithmetic mean-geometric mean inequality introduced in [1]. The result in this paper can also be viewed as a refinement of Jensen's inequality for a class of increasing and concave functions. Many other refinements can be found in [2].

Theorem 1.1 (see [1]). *Suppose that $x_k \in [a,b]$ and $p_k \geq 0$ for $k = 1,\ldots,n$, where $a > 0$, and suppose that $\sum_{k=1}^{n} p_k = 1$. Then, writing $\overline{x} = \sum_{k=1}^{n} p_k x_k$,*

$$\frac{1}{2b}\sum_{k=1}^{n} p_k \left(x_k - \overline{x}\right)^2 \leq \overline{x} - \prod_{k=1}^{n} x_k^{p_k} \leq \frac{1}{2a}\sum_{k=1}^{n} p_k \left(x_k - \overline{x}\right)^2. \tag{1.1}$$

For notational simplicity, define

$$\sigma^2 = \sum_{k=1}^{n} p_k \left(x_k - \overline{x}\right)^2 = \sum_{k=1}^{n} p_k x_k^2 - \overline{x}^2. \tag{1.2}$$

The vector $p = (p_1, \ldots, p_n)$ satisfying $p_k \geq 0$ for $k = 1, \ldots, n$ and $\sum_{k=1}^{n} p_k = 1$ will be called a *weight vector*. Sometimes, we write $\overline{x}(p)$ and $\sigma^2(p)$ to emphasize the dependency on the weight vector p. We have the following main theorem.

Theorem 1.2. *Suppose that $f(x)$ is strictly increasing, strictly concave, and twice continuously differentiable on a nonempty interval $I \subseteq \mathbb{R}$, and suppose that $f'(x)$ is strictly convex on I. Suppose that $x_k \in [a,b] \subseteq I$, where $0 < a < b$, and $p_k \geq 0$ for $k = 1, \ldots, n$, and suppose that $\sum_{k=1}^{n} p_k = 1$. Then, writing $\overline{x} = \sum_{k=1}^{n} p_k x_k$,*

(a)

$$\sum_{k=1}^{n} p_k f(x_k) \leq f(\overline{x} - \theta_1 \sigma^2), \tag{1.3}$$

where $\theta_1 = \min\{\mu_1, \mu_2\}$. Here,

$$\mu_1 = \min_{q,t,x} \frac{f'((1+qt)x) - f'((1+t)x)}{2(1-q)txf'((1+qt)x)}, \tag{1.4}$$

where the minimization is taken over $q \in [0,1]$, $t \in [0, (b-x)/x]$, and $x \in [a,b]$, and

$$\mu_2 = \min_{x \in [a,b]} -\frac{f'(x) - f'(b)}{2(x-b)f'(x)}, \tag{1.5}$$

(b)

$$\sum_{k=1}^{n} p_k f(x_k) \geq f(\overline{x} - \theta_2 \sigma^2), \tag{1.6}$$

where $\theta_2 = \max\{\pi_1, \pi_2\}$, provided $0 < \pi_1 < \infty$ and $\overline{x}(\widehat{p}) - \theta_2 \sigma^2(\widehat{p}) \in I$ for the given $x_1, \ldots, x_n$ and for all possible weight vectors $\widehat{p}$. Here,

$$\pi_1 = \left(\min_{t,q,x,\theta} \frac{2f'((1+qt)x - \theta q(1-q)t^2x^2)}{-f''((1+qt)x - \theta q(1-q)t^2x^2)} - qtx \right)^{-1}, \tag{1.7}$$

where the minimization is taken over $\theta \in [0, 1/(2(1-q)tx)]$, $q \in [0,1]$, $t \in [0, (b-x)/x]$, and $x \in [a,b]$ (we will see later that $(1+qt)x_1 - \theta q(1-q)t^2x_1^2)$ is an alternative expression for $\overline{x} - \theta\sigma^2$ when $n = 2$). And

$$\pi_2 = \max_{x \in [a,b]} \frac{f'(a) - f'(x)}{2(x-a)f'(x)}. \tag{1.8}$$

Proof. The proof is similar to the proof for Theorem 1.1 [1], based on induction on n. We first demonstrate part (a) of the theorem. The fact that the function f is always defined at $\overline{x} - \theta_1\sigma^2$ will be proved in Lemma 1.3.

The case of $n = 1$ is trivial because $\sigma^2 = 0$. When $n = 2$, write $x_2 = (1+t)x_1$, where $0 \leq t \leq (b - x_1)/x_1$, $p_1 = 1 - q$, and $p_2 = q$. With these definitions, we have $\overline{x} = (1+qt)x_1$, and $\sigma^2 = q(1-q)t^2x_1^2$. Define

$$\begin{aligned} g(t) &= f(\overline{x} - \theta_1\sigma^2) - \sum_{k=1}^{2} p_k f(x_k) \\ &= f((1+qt)x_1 - \theta_1 q(1-q)t^2x_1^2) - (1-q)f(x_1) - qf((1+t)x_1). \end{aligned} \tag{1.9}$$

Clearly, $g(0) = 0$. We will show that, for any θ_1 satisfying $0 \le \theta_1 \le \mu_1$, $g'(t) \ge 0$ for $t \ge 0$, and hence, $g(t) \ge 0$ for $t \ge 0$:

$$\begin{aligned} g'(t) &= f'\left(\overline{x} - \theta_1 \sigma^2\right)\left(qx_1 - 2\theta_1 q(1-q)tx_1^2\right) - qx_1 f'\left((1+t)x_1\right) \\ &= qx_1\left(f'\left(\overline{x} - \theta_1\sigma^2\right)\left(1 - 2\theta_1(1-q)tx_1\right) - f'\left((1+t)x_1\right)\right). \end{aligned} \tag{1.10}$$

Since $x_1 > 0$, let us ignore the factor qx_1. We wish to show, for all admissible q, t, and x_1 (Admissible parameters are those that make $x_1, \ldots, x_n$ fall on $[a, b]$. In this case, $q \in [0,1]$, $x_1 \in [a, b]$ and $t \in [0, (b - x_1)/x_1]$.),

$$f'\left(\overline{x} - \theta_1\sigma^2\right)\left(1 - 2\theta_1(1-q)tx_1\right) - f'\left((1+t)x_1\right) \ge 0. \tag{1.11}$$

Equation (1.11) is true if and only if

$$\theta_1 \le \frac{1 - f'\left((1+t)x_1\right)/f'\left(\overline{x} - \theta_1\sigma^2\right)}{2(1-q)tx_1}. \tag{1.12}$$

The right-hand side (The cases $q = 1$ or $t = 0$ do not pose problems because the right-hand side is still finite.) is an increasing function of θ_1. Substituting $\theta_1 = 0$ into the right-hand side, it suffices to show

$$\theta_1 \le \frac{f'(\overline{x}) - f'\left((1+t)x_1\right)}{2(1-q)tx_1 f'(\overline{x})} = \frac{f'\left((1+qt)x_1\right) - f'\left((1+t)x_1\right)}{2(1-q)tx_1 f'\left((1+qt)x_1\right)}. \tag{1.13}$$

Since μ_1 as in (1.4) achieves the minimum of the right-hand side above, we can choose any θ_1 satisfying $0 \le \theta_1 \le \mu_1$.

For $n > 2$, let us suppose that (1.3) has been proved for up to $n-1$, and consider the case of n. Fix $x_1, \ldots, x_n$. We may assume that all x_k are distinct. Otherwise, we can combine those identical x_k together by combining the corresponding p_k together, and we are back to the case of $n-1$ or less.

Let $p = (p_1, \ldots, p_n)$, and let

$$h(p) \triangleq f\left(\overline{x} - \theta_1\sigma^2\right) - \sum_{k=1}^{n} p_k f(x_k). \tag{1.14}$$

Define the set S by

$$S = \left\{ (p_1, \ldots, p_n) \mid p_k \ge 0, \quad \forall k \right\}. \tag{1.15}$$

We wish to show $h(p) \ge 0$ on the set S subject to the constraint $p_1 + \cdots + p_n = 1$. Suppose the minimum of $h(p)$ is in the interior of S. By the Lagrange multiplier method, any such minimum p must satisfy the following set of equations for some real number λ:

$$\frac{\partial h}{\partial p_k}(p) = \lambda \frac{\partial}{\partial p_k}\left(\sum_{k=1}^{n} p_k - 1\right) \quad \forall k. \tag{1.16}$$

This gives, for all k,

$$f'(\overline{x}-\theta_1\sigma^2)(x_k-\theta_1(x_k^2-2\overline{x}x_k))-f(x_k)=\lambda. \tag{1.17}$$

From this, we deduce that each of the x_k must satisfy the following equation with variable y:

$$f'(\overline{x}-\theta_1\sigma^2)(y-\theta_1 y^2+2\theta_1\overline{x}y)-f(y)-\lambda=0. \tag{1.18}$$

We will consider the critical points of the left-hand side above, that is, the zeros of its derivative with respect to y. By taking the derivative, the critical points are the solutions to the equation

$$f'(\overline{x}-\theta_1\sigma^2)(2\theta_1(\overline{x}-y)+1)=f'(y). \tag{1.19}$$

The left-hand side of (1.19) is a decreasing linear function of y. Under the assumption that $f'(y)$ is strictly convex, there can be at most two solutions to the equation. We will show that, under suitable conditions, there is at most one solution. The situation is illustrated in Figure 1. We will find conditions for the following to hold, for any admissible $\overline{x}$ and σ^2,

$$f'(\overline{x}-\theta_1\sigma^2)(2\theta_1(\overline{x}-b)+1)>f'(b). \tag{1.20}$$

When the x_k are not all identical, $\sigma^2\neq 0$. Hence, for $\theta_1>0$, $f'(\overline{x}-\theta_1\sigma^2)>f'(\overline{x})>0$, it is enough to consider

$$f'(\overline{x})(2\theta_1(\overline{x}-b)+1)\geq f'(b) \tag{1.21}$$

which is the same as

$$\theta_1\leq-\frac{f'(\overline{x})-f'(b)}{2(\overline{x}-b)f'(\overline{x})}. \tag{1.22}$$

We can choose θ_1 as in the theorem, which satisfies $\theta_1\leq\mu_1$ and

$$\theta_1\leq\mu_2=\min_{x\in[a,b]}-\frac{f'(x)-f'(b)}{2(x-b)f'(x)}. \tag{1.23}$$

By Rolle's theorem, we conclude that there can be at most two distinct roots to (1.18) on the interval $[a,b]$. This contradicts our assumption that all $x_1,\ldots,x_n$ are distinct. Hence, it must be true that the minimum of $h(p)$ in S subject to the constraint $p_1+\cdots+p_n=1$ occurs on the boundary of S, where some of the p_k must be zero. At the minimum, say p^o, we are back to the case of $n-1$ or less, and by the induction hypothesis, $h(p^o)\geq 0$. Hence, $h(p)\geq 0$ for arbitrary $p\in S$ subject to the constraint $p_1+\cdots+p_n=1$.

We now proceed to show (1.6) in part (b). For the case of $n=2$, let us replace θ_1 by θ_2 in (1.9), and rename the function $\widehat{g}(t)$. Again $\widehat{g}(0)=0$. We will find appropriate θ_2 for which $\widehat{g}'(t)\leq 0$ for $t\geq 0$. Following similar steps as before, we get

$$\widehat{g}'(t)=qx_1(f'(\overline{x}-\theta_2\sigma^2)(1-2\theta_2(1-q)tx_1)-f'((1+t)x_1)). \tag{1.24}$$

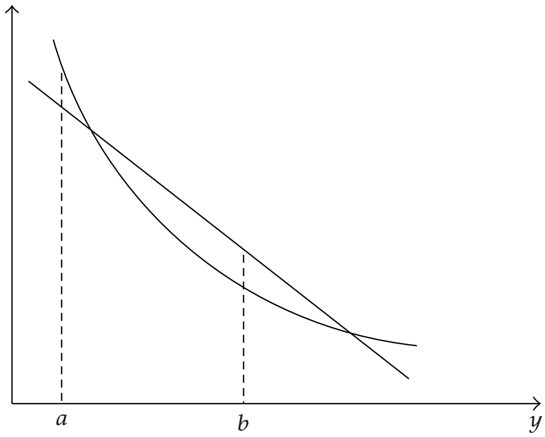

Figure 1: Illustration of functions $f'(\overline{x}-\theta_1\sigma^2)(2\theta_1(\overline{x}-y)+1)$ and $f'(y)$.

To show $\widehat{g}'(t) \le 0$, it suffices to show

$$f'(\overline{x}-\theta_2\sigma^2)\big(1-2\theta_2(1-q)tx_1\big)-f'\big((1+t)x_1\big) \le 0. \tag{1.25}$$

Observe that if $1-2\theta_2(1-q)tx_1 \le 0$, or equivalently, $\theta_2 \ge 1/(2(1-q)tx_1)$, the above automatically holds. (We assume the convention $1/0=\infty$.) We sketch our subsequent strategy. For any fixed q, t, and x_1, find $\theta_2(q,t,x_1)$ that satisfies both (1.25) and

$$\theta_2(q,t,x_1) \le \frac{1}{2(1-q)tx_1}. \tag{1.26}$$

Such $\theta_2(q,t,x_1)$ must exist because $1/(2(1-q)tx_1)$ qualifies. Then, we can take $\sup\theta_2(q,t,x_1)$ over all admissible q, t, and x_1.

By the mean value theorem, there exists $\widehat{\eta} \in (\overline{x}-\theta_2\sigma^2,(1+t)x_1)$ such that

$$\begin{aligned} &f'(\overline{x}-\theta_2\sigma^2)\big(1-2\theta_2(1-q)tx_1\big)-f'\big((1+t)x_1\big) \\ &\quad= -f''(\widehat{\eta})(1-q)tx_1\big(1+\theta_2 qtx_1\big)-f'(\overline{x}-\theta_2\sigma^2)2\theta_2(1-q)tx_1. \end{aligned} \tag{1.27}$$

Note that, for (1.25) to hold, it suffices if

$$-f''(\widehat{\eta})\big(1+\theta_2 qtx_1\big)-f'(\overline{x}-\theta_2\sigma^2)2\theta_2 \le 0. \tag{1.28}$$

Since $-f''(x)$ is decreasing and positive, (1.28) is implied by

$$-f''(\overline{x}-\theta_2\sigma^2)\big(1+\theta_2 qtx_1\big)-f'(\overline{x}-\theta_2\sigma^2)2\theta_2 \le 0, \tag{1.29}$$

which is equivalent to

$$\frac{1}{\theta_2} \le \frac{2f'(\overline{x}-\theta_2\sigma^2)}{-f''(\overline{x}-\theta_2\sigma^2)}-qtx_1. \tag{1.30}$$

We wish to find θ_2 that satisfies (1.30) for all admissible q, t, and x_1.

For any fixed q, t, and x_1, suppose we obtain

$$\theta_2(q,t,x_1) = \left(\min_{\theta\in[0,1/(2(1-q)tx_1)]} \frac{2f'(\overline{x}-\theta\sigma^2)}{-f''(\overline{x}-\theta\sigma^2)} - qtx_1\right)^{-1} \tag{1.31}$$

and suppose $0 < \theta_2(q,t,x_1) < \infty$. Let us consider $\widehat{\theta}_2 \geq \theta_2(q,t,x_1)$. If $\widehat{\theta}_2 > 1/(2(1-q)tx_1)$, then (1.25) holds trivially. If $\widehat{\theta}_2 \leq 1/(2(1-q)tx_1)$, then

$$\frac{1}{\widehat{\theta}_2} \leq \frac{1}{\theta_2(q,t,x_1)} = \min_{\theta\in[0,1/(2(1-q)tx_1)]} \frac{2f'(\overline{x}-\theta\sigma^2)}{-f''(\overline{x}-\theta\sigma^2)} - qtx_1 \leq \frac{2f'(\overline{x}-\theta_2\sigma^2)}{-f''(\overline{x}-\theta_2\sigma^2)} - qtx_1, \tag{1.32}$$

that is, (1.30) holds, which implies that (1.28), and hence, (1.25) hold. Hence, we can choose $\theta_2 \geq \sup_{q,t,x_1}\theta_2(q,t,x_1)$, where the supremum is over all admissible q, t, and x_1. To summarize, we can choose the following θ_2 for the theorem, provided $0 < \pi_1 < \infty$,

$$\theta_2 \geq \pi_1 = \left(\min_{t,q,x_1,\theta} \frac{2f'(\overline{x}-\theta\sigma^2)}{-f''(\overline{x}-\theta\sigma^2)} - qtx_1\right)^{-1}, \tag{1.33}$$

where the minimization is taken over $\theta \in [0,1/(2(1-q)tx_1)]$, and over all admissible q, x_1, t, that is, $q \in [0,1]$, $x_1 \in [a,b]$, and $t \in [0,(b-x_1)/x_1]$. This minimization can be solved easily in a number of cases, which we will show later.

For $n > 2$, the proof is nearly identical to that for part (a). Let us suppose (1.6) has been proved for up to $n-1$, and consider the case of n. Fix $x_1,\ldots,x_n$. We may assume that all x_k are distinct as before. Let

$$\widehat{h}(p) \triangleq f(\overline{x}-\theta_2\sigma^2) - \sum_{k=1}^{n} p_k f(x_k). \tag{1.34}$$

We wish to show $\widehat{h}(p) \leq 0$ on the set S defined before, subject to the constraint $p_1+\cdots+p_n = 1$. Suppose the maximum of $\widehat{h}(p)$ is in the interior of S. Applying the Lagrange multiplier method for finding the constrained maximum of $\widehat{h}(p)$ on S, we deduce that any maximum p must satisfy the following set of equations for some real number λ:

$$\frac{\partial\widehat{h}}{\partial p_k}(p) = \lambda\frac{\partial}{\partial p_k}\left(\sum_{k=1}^{n} p_k - 1\right) \quad \forall k. \tag{1.35}$$

This gives, for all k,

$$f'(\overline{x}-\theta_2\sigma^2)(x_k - \theta_2(x_k^2 - 2\overline{x}x_k)) - f(x_k) = \lambda. \tag{1.36}$$

Each of the x_k must satisfy the following equation with variable y:

$$f'(\overline{x}-\theta_2\sigma^2)(y - \theta_2 y^2 + 2\theta_2\overline{x}y) - f(y) - \lambda = 0. \tag{1.37}$$

We will consider the critical points of the left-hand side above, which are the solutions to the equation

$$f'(\overline{x}-\theta_2\sigma^2)(2\theta_2(\overline{x}-y)+1) = f'(y). \tag{1.38}$$

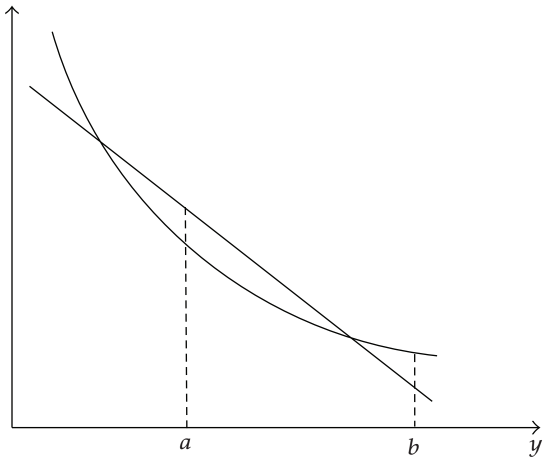

Figure 2: Illustration of functions $f'(\overline{x} - \theta_2\sigma^2)(2\theta_2(\overline{x} - y) + 1)$ and $f'(y)$.

Under the assumption that $f'(y)$ is strictly convex, there can be at most two solutions to the equation. We will show that, under suitable conditions, there is at most one solution. The situation is illustrated in Figure 2. We will find conditions for the following to hold, for any admissible $\overline{x}$ and σ^2:

$$f'\left(\overline{x} - \theta_2\sigma^2\right)\left(2\theta_2\left(\overline{x} - a\right) + 1\right) > f'(a). \tag{1.39}$$

When the x_k are not all identical, $\sigma^2 \neq 0$. Hence, for $\theta_2 > 0$, $f'(\overline{x} - \theta_2\sigma^2) > f'(\overline{x}) > 0$, it is enough to consider

$$f'\left(\overline{x}\right)\left(2\theta_2\left(\overline{x} - a\right) + 1\right) \geq f'(a) \tag{1.40}$$

which is the same as

$$\theta_2 \geq -\frac{f'\left(\overline{x}\right) - f'(a)}{2\left(\overline{x} - a\right)f'\left(\overline{x}\right)}. \tag{1.41}$$

We can choose θ_2 as in the theorem, which satisfies $\theta_2 \geq \pi_1$ and

$$\theta_2 \geq \pi_2 = \max_{x \in [a,b]} -\frac{f'(x) - f'(a)}{2(x - a)f'(x)}. \tag{1.42}$$

By Rolle's theorem, we conclude that there can be at most two distinct roots to (1.37) on the interval $[a, b]$. This contradicts our assumption that all $x_1, \ldots, x_n$ are distinct. Hence, it must be true that the maximum of $\widehat{h}(p)$ in S subject to the constraint $p_1 + \cdots + p_n = 1$ occurs on the boundary of S, where some of the p_k must be zero. At the maximum, say p^o, we are back to the case of $n - 1$ or less, and by the induction hypothesis, $\widehat{h}(p^o) \leq 0$. Hence, $\widehat{h}(p) \leq 0$ for arbitrary $p \in S$ subject to the constraint $p_1 + \cdots + p_n = 1$. □

We now complete the proof for part (a) of Theorem 1.2 by showing the following lemma.

Lemma 1.3. *For any integer $n \geq 1$, and for all $x_1, \ldots, x_n$ with each $x_k \in [a, b]$ and all $p_1, \ldots, p_n$, where each $p_k \geq 0$ and $\sum_{k=1}^{n} p_k = 1$,*

$$\overline{x} - \theta_1 \sigma^2 \geq a, \tag{1.43}$$

where θ_1 is as given in Theorem 1.2.

Proof. It suffices to show

$$\overline{x} - \mu_1 \sigma^2 \geq a. \tag{1.44}$$

The case of $n = 1$ is trivial, since $\sigma^2 = 0$. For the case $n = 2$, as before, write $x_2 = (1 + t)x_1$, where $0 \leq t \leq (b - x_1)/x_1$, $p_1 = 1 - q$, and $p_2 = q$. With these, we have $\overline{x} = (1 + qt)x_1$, and $\sigma^2 = q(1 - q)t^2 x_1^2$. In the cases where $q = 0$, $q = 1$ or $t = 0$, (1.44) holds trivially. For $0 < q < 1$ and $t > 0$, (1.44) holds if and only if

$$\mu_1 \leq \psi(q) \triangleq \frac{(1 + qt)x_1 - a}{q(1 - q)t^2 x_1^2}. \tag{1.45}$$

We will show that this is indeed true for all admissible q, t, and x_1. For the case $x_1 = a$, $\psi(q) = 1/((1-q)ta)$. Because $ta \leq b-a$, $\psi(q) \geq 1/(b-a)$. When $a < x_1 \leq b$, the value of $\psi(q)$ approaches $+\infty$ as q approaches 0 or 1. The minimum value must be on $(0, 1)$. The derivative of $\psi(q)$ is

$$\psi'(q) = \frac{q(1 - q)tx_1 + ((1 + qt)x_1 - a)(2q - 1)}{(1 - q)^2 q^2 t^2 x_1^2}. \tag{1.46}$$

For q_o that satisfies $\psi'(q_o) = 0$, we have the following identity:

$$(1 + q_o t)x_1 - a = \frac{q_o(1 - q_o)tx_1}{1 - 2q_o}. \tag{1.47}$$

Hence,

$$\psi(q_o) = \frac{1}{(1 - 2q_o)tx_1}. \tag{1.48}$$

Because $tx_1 \leq b - a$, we get $\psi(q_o) \geq 1/(b - a)$. By the definition of μ_1, for all admissible q, t, and x_1,

$$\mu_1 \leq \frac{f'(\overline{x}) - f'((1 + t)x_1)}{2(1 - q)tx_1 f'(\overline{x})} \leq \frac{1}{2(1 - q)tx_1}. \tag{1.49}$$

Hence, $\mu_1 \leq 1/(2(b - a))$. Therefore, (1.45) holds for all admissible q, t, and x_1.

Consider the case of $n \geq 3$. Fix $x_1, \ldots, x_n$ and suppose they are all distinct. Let

$$\phi(p) = \overline{x} - \mu_1 \sigma^2 - a = \sum_{k=1}^{n} p_k x_k - \mu_1 \sum_{k=1}^{n} p_k x_k^2 + \mu_1 \left(\sum_{k=1}^{n} p_k x_k \right)^2 - a. \tag{1.50}$$

Consider minimizing $\phi(p)$ over all possible weight vectors and suppose p^o is the minimum. Then, there exists a constant λ such that, for any k with $p_k^o > 0$, we must have $(\partial\phi/\partial p_k)(p^o) = \lambda$. That is,

$$x_k - \mu_1 (x_k^2 - 2\overline{x}(p^o)x_k) = \lambda, \quad \forall k \text{ with } p_k^o > 0. \tag{1.51}$$

For the given p^o, there can be at most two distinct x_k satisfying the equation $y - \mu_1(y^2 - 2\overline{x}(p^o)y) = \lambda$ in variable y. Hence, there can be at most two nonzero components in p^o. This belongs to the $n = 2$ case and $\phi(p^o) \geq 0$. Therefore, for all weight vectors p, $\phi(p) \geq 0$. □

For part (b) of Theorem 1.2, the proof requires $\overline{x}(p) - \theta_2\sigma^2(p)$ to be within the domain of the function f for various unknown weight vectors p. This is why the statement of Theorem 1.2 makes the assumption that this is true for all possible weight vectors. The following lemma gives a simple sufficient condition for this assumption to hold.

Lemma 1.4. *Fix $x_1, \dots, x_n$ on $[a,b]$, $n \geq 2$. Let θ_2 be as given in Theorem 1.2. Without loss of generality, assume $x_1 < x_2 < \cdots < x_n$. Then,*

(a) when $\theta_2 \leq 1/(x_n - x_1)$,

$$\overline{x}(p) - \theta_2\sigma^2(p) \geq x_1 \quad \text{for all weight vectors } p, \tag{1.52}$$

and hence, $\overline{x}(p) - \theta_2\sigma^2(p) \in [a,b]$ for all weight vectors p;

(b) when $\theta_2 > 1/(x_n - x_1)$,

$$\overline{x}(p) - \theta_2\sigma^2(p) \geq x_1 - \frac{\left(\theta_2(x_n - x_1) - 1\right)^2}{4\theta_2} \quad \text{for all weight vectors } p. \tag{1.53}$$

Hence, if $[\min\{a, x_1 - (\theta_2(x_n - x_1) - 1)^2/(4\theta_2)\}, b] \subseteq I$, then $\overline{x}(p) - \theta_2\sigma^2(p) \in I$ for all weight vectors p.

Proof. Write $p = (p_1, \dots, p_n)$. Consider the minimization problem,

$$\min_{p \geq 0,\ \sum_{k=1}^{n} p_k = 1} \overline{x}(p) - \theta_2\sigma^2(p). \tag{1.54}$$

Using the same argument as in the proof of Lemma 1.3, we can conclude that the minimum is achieved at some p with at most two nonzero components. Hence, it suffices to consider minimization problems of the following form:

$$\min_{p_i \geq 0,\ p_j \geq 0,\ p_i + p_j = 1} p_i x_i + p_j x_j - \theta_2\left(p_i x_i^2 + p_j x_j^2 - \left(p_i x_i + p_j x_j\right)^2\right). \tag{1.55}$$

It remains to be decided which i and j should be used in the above minimization. Suppose $x_i < x_j$ (here, i and j are unknown indices). We claim that $i = 1$. To see this, the partial derivative of the objective function with respect to x_i is $p_i - \theta_2(2p_i x_i - 2\overline{x}p_i)$, where $\overline{x} = p_i x_i + p_j x_j$. Since $x_i \leq \overline{x}$, the partial derivative is nonnegative, and hence, the function is nondecreasing in x_i.

Once x_i is chosen to be x_1, the second partial derivative of the function with respect to x_j is $-2\theta_2(p_j - p_j^2)$, which is nonpositive. The minimum is achieved at either $x_j = x_1$ or $x_j = x_n$.

To summarize, the original minimization problem (1.54) is achieved either at $p_1 = 1$, in which case the minimum value is x_1, or it has the same minimum value as the following problem:

$$\min_{p_1 \geq 0, p_n \geq 0, p_1 + p_n = 1} p_1 x_1 + p_n x_n - \theta_2\left(p_1 x_1^2 + p_n x_n^2 - \left(p_1 x_1 + p_n x_n\right)^2\right). \tag{1.56}$$

It is easy to show that, if $\theta_2 \leq 1/(x_n - x_1)$, the minimum of (1.60) is achieved at $p_1 = 1$ and the minimum value is x_1. Otherwise, the minimum is achieved at

$$p_1 = \frac{1}{2}\left(1 + \frac{1}{\theta_2(x_n - x_1)}\right), \qquad p_2 = \frac{1}{2}\left(1 - \frac{1}{\theta_2(x_n - x_1)}\right), \tag{1.57}$$

and the minimum value is $x_1 - (\theta_2(x_n - x_1) - 1)^2/(4\theta_2)$, which is no greater than x_1 for all $\theta_2 > 0$. □

We now make some remarks about Theorem 1.2.

Remark 1.5. For part (a) of the theorem, we can chose a smaller value, $-u/(2f'(a))$, for μ_1 than that in (1.4). Let $u \le 0$ be an upper bound of $f''(x)$ on $[a,b]$. Note that

$$(1+qt)x_1 - (1+t)x_1 = -(1-q)tx_1. \tag{1.58}$$

By the mean value theorem, there exists some $\eta \in ((1+qt)x_1, (1+t)x_1)$ such that

$$\frac{f'((1+qt)x_1) - f'((1+t)x_1)}{2(1-q)tx_1 f'((1+qt)x_1)} = \frac{-f''(\eta)}{2f'((1+qt)x_1)} \ge \frac{-u}{2f'(a)}. \tag{1.59}$$

Therefore, we can choose $-u/(2f'(a))$ for μ_1.

Remark 1.6. For part (b) of the theorem, two simpler but less widely applicable choices for θ_2 can be deduced from (1.33). Since $qtx_1 \le b-a$, π_1 can be relaxed to

$$\pi_1^1 = \frac{1}{\min_{x\in[a,b]} 2f'(x)/(-f''(x)) - (b-a)}. \tag{1.60}$$

We must require $\min_{x\in[a,b]} 2f'(x)/(-f''(x)) - (b-a) > 0$ for π_1^1 to be useful. An even simpler choice is, when $2f'(b) + f''(a)(b-a) > 0$,

$$\pi_1^2 = \frac{1}{2f'(b)/(-f''(a)) - (b-a)} = \frac{-f''(a)}{2f'(b) + f''(a)(b-a)}. \tag{1.61}$$

Note that $0 < \pi_1^2 < \infty$ implies $0 < \pi_1^1 < \infty$, which, in turn, implies $0 < \pi_1 < \infty$. In this case, $\pi_1 \le \pi_1^1 \le \pi_1^2$. Similarly, if $0 < \pi_1^1 < \infty$, then $\pi_1 \le \pi_1^1$.

π_2 as in (1.8) can also be relaxed. Since, for some $\xi \in (a,b)$,

$$-\frac{f'(x) - f'(a)}{2(x-a)f'(x)} = -\frac{f''(\xi)}{2f'(x)}, \tag{1.62}$$

a relaxation is

$$\pi_2^1 = \frac{-f''(a)}{2f'(b)}. \tag{1.63}$$

Hence, for part (b) of the theorem, we can choose $\theta_2 = \max\{\pi_1^1, \pi_2^1\}$, provided $0 < \pi_1^1 < \infty$. Alternatively, we can choose $\theta_2 = \max\{\pi_1^2, \pi_2^1\} = \pi_1^2$, provided $0 < \pi_1^2 < \infty$.

Remark 1.7. The strength of Theorem 1.2 is in bringing the variations of x_k's into consideration, through σ^2. There are alternative methods for finding tighter upper bound of $\sum_{k=1}^{n} p_k f(x_k)$ than that given by Jensen's inequality, $\sum_{k=1}^{n} p_k f(x_k) \le f(\overline{x})$. For instance, we can apply the arithmetic mean-geometric mean inequality. For any α,

$$\prod_{k=1}^{n} \left(e^{\alpha f(x_k)}\right)^{p_k} \le \sum_{k=1}^{n} p_k e^{\alpha f(x_k)}. \tag{1.64}$$

Hence, for any $\alpha > 0$,

$$\sum_{k=1}^{n} p_k f(x_k) \le \frac{1}{\alpha} \log\left(\sum_{k=1}^{n} p_k e^{\alpha f(x_k)}\right). \tag{1.65}$$

Then, choose a small positive α.

2. Special cases and examples

The following theorem is needed.

Theorem 2.1 (see [3, page 4]). *Let I be a nonempty interval of $\mathbb{R}$. A function $h : I \to \mathbb{R}$ is convex if and only if, for all $x_o \in I$, $(h(x) - h(x_o))/(x - x_o)$ is a nondecreasing function of x on $I \setminus \{x_o\}$.*

Corollary 2.2. *Let I be a nonempty interval of $\mathbb{R}$. Suppose the function $h : I \to \mathbb{R}$ is convex, and $h(x) \neq 0$ for all $x \in I$. If $1/h(x)$ is a convex function, then, for all $x_o \in I$, $(h(x)-h(x_o))/((x-x_o)h(x))$ is a nonincreasing function of x on $I \setminus \{x_o\}$. If $1/h(x)$ is a concave function, then, for all $x_o \in I$, $(h(x) - h(x_o))/((x - x_o)h(x))$ is a nondecreasing function of x on $I \setminus \{x_o\}$.*

Proof. Consider

$$\frac{h(x) - h(x_o)}{(x - x_o)h(x)} = \frac{1 - h(x_o)/h(x)}{(x - x_o)} = -\frac{g(x) - g(x_o)}{x - x_o}, \tag{2.1}$$

where we define $g(x) = h(x_o)/h(x)$. If $1/h(x)$ is convex, then $g(x)$ is convex. By Theorem 2.1, $(h(x) - h(x_o))/((x - x_o)h(x))$ is nonincreasing. If $1/h(x)$ is concave, then $g(x)$ is concave, and hence, $(h(x) - h(x_o))/((x - x_o)h(x))$ is nondecreasing. □

Theorem 2.3. *Consider the function $f(x)$ as specified in Theorem 1.2. In addition, suppose $f'(x)$ is convex. If $1/f'(x)$ is concave, then*

$$\theta_1 = \min_{x\in[a,b]} \frac{-f''(x)}{2f'(x)}. \tag{2.2}$$

If $1/f'(x)$ is convex, then

$$\theta_1 = \frac{f'(a) - f'(b)}{2(b - a)f'(a)}. \tag{2.3}$$

Proof. If $1/f'(x)$ is a concave function, then $1/f'((1 + qt)x_1)$ is a concave function in q for $q \in [0, 1]$. By Corollary 2.2, $(f'((1+qt)x_1)-f'((1+t)x_1))/(2(1-q)tx_1 f'((1+qt)x_1))$ is nonincreasing in q. Its minimum with respect to q occurs at $q = 1$. Substituting $q = 1$, we get

$$\mu_1 = \min_{t,x_1} \frac{-f''((1 + t)x_1)}{2f'((1 + t)x_1)} = \min_{q,t,x_1} \frac{f'((1 + qt)x_1) - f'((1 + t)x_1)}{2(1 - q)tx_1 f'((1 + qt)x_1)} = \min_{x\in[a,b]} \frac{-f''(x)}{2f'(x)}. \tag{2.4}$$

On the other hand, by Corollary 2.2, $-(f'(x)-f'(b))/(2(x-b)f'(x))$ is a nonincreasing function of x. Hence,

$$\mu_2 = -\frac{f''(b)}{2f'(b)} \geq \mu_1. \tag{2.5}$$

If $1/f'(x)$ is a convex function, $(f'((1 + qt)x_1) - f'((1 + t)x_1))/(2(1 - q)tx_1 f'((1 + qt)x_1))$ achieves its minimum with respect to q at $q = 0$. Substituting $q = 0$, we get

$$\mu_1 = \min_{t,x_1} \frac{f'(x_1) - f'((1 + t)x_1)}{2tx_1 f'(x_1)}. \tag{2.6}$$

By Theorem 2.1, $(f'(x_1) - f'((1+t)x_1))/(2tx_1)$ is a nonincreasing function of t, and hence, its minimum occurs at $t = (b - x_1)/x_1$. Hence,

$$\mu_1 = \min_{x_1 \in [a,b]} \frac{f'(x_1) - f'(b)}{2(b - x_1) f'(x_1)} = \mu_2. \tag{2.7}$$

By Corollary 2.2, $(f'(x_1) - f'(b))/(2(b - x_1)f'(x_1))$ is a nondecreasing function of x_1. Its minimum, which occurs at $x_1 = a$, is $(f'(a) - f'(b))/(2(b-a)f'(a))$. □

Example 2.4 ($f(x) = \log x$). Here, $f'(x) = 1/x$ and $f''(x) = -1/x^2$. Hence, $f'(x)$ and $1/f'(x)$ are both convex functions. We can choose

$$\begin{aligned}
\theta_1 &= \min_{q,t,x_1} \frac{f'((1+qt)x_1) - f'((1+t)x_1)}{2(1-q)tx_1 f'((1+qt)x_1)} = \min_{t,x_1} \frac{1}{2(1+t)x_1} = \frac{1}{2b}, \\
\theta_2(q,t,x_1) &= \left(\min_{\theta \in [0,1/(2(1-q)tx_1)]} \frac{2f'(\overline{x} - \theta\sigma^2)}{-f''(\overline{x} - \theta\sigma^2)} - qtx_1 \right)^{-1} \\
&= \left(\min_{\theta \in [0,1/(2(1-q)tx_1)]} 2(\overline{x} - \theta\sigma^2) - qtx_1 \right)^{-1} \\
&= \left(2\left(\overline{x} - \frac{1}{2(1-q)tx_1}\sigma^2 \right) - qtx_1 \right)^{-1} \\
&= \left(2\left(x_1 + qtx_1 - \frac{q(1-q)t^2x_1^2}{2(1-q)tx_1} \right) - qtx_1 \right)^{-1} \\
&= (2x_1)^{-1}.
\end{aligned} \tag{2.8}$$

Maximizing over x_1, we get $\pi_1 = 1/(2a)$. Also,

$$\pi_2 = \max_{x \in [a,b]} -\frac{f'(x) - f'(a)}{2(x-a)f'(x)} = \frac{1}{2a}. \tag{2.9}$$

Example 2.5 ($f(x) = x^\alpha$, where $0 < \alpha < 1$). Here, $f'(x) = \alpha x^{\alpha-1}$ is convex, and $f''(x) = \alpha(\alpha - 1)x^{\alpha-2}$. Because $1/x^{\alpha-1}$ is a concave function, we can apply Theorem 2.3, and get

$$\begin{aligned}
\theta_1 &= \min_{x \in [a,b]} \frac{-f''(x)}{2f'(x)} = \min_{x \in [a,b]} \frac{1-\alpha}{2x} = \frac{1-\alpha}{2b}, \\
\theta_2(q,t,x_1) &= \left(\min_{\theta \in [0,1/(2(1-q)tx_1)]} \frac{2f'(\overline{x} - \theta\sigma^2)}{-f''(\overline{x} - \theta\sigma^2)} - qtx_1 \right)^{-1} \\
&= \left(\min_{\theta \in [0,1/(2(1-q)tx_1)]} \frac{2(\overline{x} - \theta\sigma^2)}{1-\alpha} - qtx_1 \right)^{-1} \\
&= \left(\frac{2x_1 + qtx_1}{1-\alpha} - qtx_1 \right)^{-1} \\
&= \left(\frac{2x_1}{1-\alpha} + \frac{\alpha}{1-\alpha} qtx_1 \right)^{-1}.
\end{aligned}$$

Maximizing the last expression over q, we get $(1-\alpha)/(2x_1)$. Maximizing the result over x_1, we get

$$\pi_1 = \frac{1-\alpha}{2a}. \tag{2.10}$$

Also, by Corollary 2.2, $-(f'(x)-f'(a))/(2(x-a)f'(x))$ is a nonincreasing function of x, and achieves its maximum at $x=a$. Hence,

$$\pi_2 = \max_{x\in[a,b]} -\frac{f'(x)-f'(a)}{2(x-a)f'(x)} = -\frac{f''(a)}{2f'(a)} = \frac{1-\alpha}{2a}. \tag{2.11}$$

Example 2.6 ($f(x) = -e^{-x}$). Here, $f'(x) = e^{-x}$ and $f''(x) = -e^{-x}$. Because $1/f'(x)$ is convex, we can apply Theorem 2.3, and get

$$\theta_1 = \frac{f'(a)-f'(b)}{2(b-a)f'(a)} = \frac{1-e^{-(b-a)}}{2(b-a)},$$

$$\theta_2(q,t,x_1) = \left(\min_{\theta\in[0,1/(2(1-q)tx_1)]} \frac{2f'(\overline{x}-\theta\sigma^2)}{-f''(\overline{x}-\theta\sigma^2)} - qtx_1\right)^{-1} = (2-qtx_1)^{-1}.$$

Maximizing the last expression over qtx_1, we get

$$\pi_1 = \frac{1}{2-(b-a)}. \tag{2.12}$$

We must require $b-a<2$. Also, by Corollary 2.2, $-(f'(x)-f'(a))/(2(x-a)f'(x))$ is a nondecreasing function of x, and achieves its maximum at $x=b$. Hence,

$$\pi_2 = \max_{x\in[a,b]} -\frac{f'(x)-f'(a)}{2(x-a)f'(x)} = -\frac{f'(b)-f'(a)}{2(b-a)f'(b)} = \frac{e^{b-a}-1}{2(b-a)}. \tag{2.13}$$

Lemma 2.7. *For* $0<b-a<2$,

$$\frac{1}{2-(b-a)} \geq \frac{e^{b-a}-1}{2(b-a)}. \tag{2.14}$$

Proof. Note that, for $0<b-a<2$, (2.14) is equivalent to

$$2(b-a) \geq 2(e^{b-a}-1) - (b-a)e^{b-a} + (b-a). \tag{2.15}$$

Let $v=b-a$, and assume $v\in[0,2]$. It suffices to show

$$v - 2(e^v-1) + ve^v \geq 0. \tag{2.16}$$

Since at $v=0$, the left hand above is 0. We need only to show that $v-2(e^v-1)+ve^v$ is nondecreasing function of v for $v\geq 0$. Its derivative is $1-e^v+ve^v$, which is equal to zero at $v=0$. But, for $v\geq 0$,

$$(1-e^v+ve^v)' = ve^v \geq 0. \tag{2.17}$$

Hence, $1-e^v+ve^v\geq 0$ for $v\geq 0$, and (2.16) is true. $\square$

Example 2.8 ($f(x) = -1/x^k$ where $k > 0$). Here, $f'(x) = k/x^{k+1}$, $f''(x) = -k(k+1)/x^{k+1}$. Since $1/f'(x) = x^{k+1}/k$ is convex, by Theorem 2.3,

$$\theta_1 = \frac{f'(a) - f'(b)}{2(b-a)f'(a)} = \frac{1 - (a/b)^{k+1}}{2(b-a)},$$

$$\begin{aligned}
\theta_2(q,t,x_1) &= \left(\min_{\theta\in[0,1/(2(1-q)tx_1)]} \frac{2f'(\overline{x} - \theta\sigma^2)}{-f''(\overline{x} - \theta\sigma^2)} - qtx_1\right)^{-1} \\
&= \left(\min_{\theta\in[0,1/(2(1-q)tx_1)]} \frac{2(\overline{x} - \theta\sigma^2)}{k+1} - qtx_1\right)^{-1} \\
&= \left(\frac{2x_1 + qtx_1}{k+1} - qtx_1\right)^{-1} \\
&= \left(\frac{2x_1}{k+1} - \frac{k}{k+1}qtx_1\right)^{-1}.
\end{aligned}$$

Maximizing the last expression over q and t, we get $1/((k+2)x_1/(k+1) - bk/(k+1))$. Maximizing the result over x_1, we get

$$\pi_1 = \frac{k+1}{(k+2)a - bk}. \tag{2.18}$$

We must require $(k+2)a - bk > 0$, or $k < 2a/(b-a)$, for the above π_1 to be useful for Theorem 1.2. Also, by Corollary 2.2, $-(f'(x) - f'(a))/(2(x-a)f'(x))$ is a nondecreasing function of x, and achieves its maximum at $x = b$:

$$\pi_2 = \max_{x\in[a,b]} -\frac{f'(x) - f'(a)}{2(x-a)f'(x)} = -\frac{f'(b) - f'(a)}{2(b-a)f'(b)} = \frac{(b/a)^{k+1} - 1}{2(b-a)}. \tag{2.19}$$

References

[1] D. I. Cartwright and M. J. Field, "A refinement of the arithmetic mean-geometric mean inequality," *Proceedings of the American Mathematical Society*, vol. 71, no. 1, pp. 36–38, 1978.

[2] J. E. Pečarić, F. Proschan, and Y. L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*, vol. 187 of *Mathematics in Science and Engineering*, Academic Press, Boston, Mass, USA, 1992.

[3] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms. I: Fundamentals*, vol. 305 of *Grundlehren der Mathematischen Wissenschaften*, Springer, Berlin, Germany, 1993.