CONSTRUCTING GENERALIZED MEAN FUNCTIONS USING CONVEX FUNCTIONS WITH REGULARITY CONDITIONS

YUN-BIN ZHAO*, SHU-CHERNG FANG†, AND DUAN LI ‡

Abstract. The generalized mean function has been widely used in convex analysis and mathematical programming. This paper studies a further generalization of such a function. A necessary and sufficient condition is obtained for the convexity of a generalized function. Additional sufficient conditions that can be easily checked are derived for the purpose of identifying some classes of functions which guarantee the convexity of the generalized functions. We show that some new classes of convex functions with certain regularity (such as S^*-regularity) can be used as building blocks to construct such generalized functions.

Key words. Convexity, mathematical programming, generalized mean function, self-concordant functions, S^*-regular functions.

AMS subject classifications. 90C30, 90C25, 52A41, 49J52

1. Introduction. In this paper, we denote the n-dimensional Euclidean space by R^n, its nonnegative orthant by R^n_+, and positive orthant by R^n_{++}.

In 1934, Hardy, Littlewood and Pólya ([13]) considered the following function under the name of generalized mean:

$$\Upsilon_w(x) = \phi^{-1}\left(\sum_{i=1}^n w_i \phi(x_i)\right) \tag{1.1}$$

where $\phi(\cdot)$ is a real, strictly increasing, convex function defined on a subset of R and $w = (w_1, w_2, \ldots, w_n)^T$ is a given vector in R^n_+. Assuming that $\phi > 0$, $\phi' > 0$ and $\phi'' > 0$, they showed an equivalent condition for the convexity of Υ_w. When ϕ is three times differentiable, Ben-Tal and Teboulle ([2]) established another equivalent condition for Υ_w being convex (see next section for details).

The generalized mean function (1.1) has many applications in optimization. Ben-Tal and Teboulie ([2]) demonstrated an interesting application of (1.1) (in a continuous form) on penalty functions and duality formulation of stochastic nonlinear programming problems. However, the most widely used generalized means are the logorithmic-exponentional and p-norm functions:

$$f_w(x) = \log\left(\sum_{i=1}^n w_i e^{x_i}\right), \quad p_w(x) = \left(\sum_{i=1}^n w_i x_i^p\right)^{1/p} \quad \text{for } x = (x_1, ..., x_n)^T \in R^n.$$

They correspond to the special cases of Υ_w with $\phi(t) = e^t$ and $\phi(t) = t^p$, respectively.

Needless to say that the log-exp function has been widely used in convex analysis and mathematical programming. For example, a geometric program (see Duffin *et*

*Institute of Applied Mathematics, AMSS, Chinese Academy of Sciences, Beijing 100080, China (Email: ybzhao@amss.ac.cn). This author's work was partially supported by the National Natural Science Foundation of China under Grant No. 10201032 and Grant No. 70221001.

†Corresponding author. Industrial Engineering and Operations Research, North Carolina State University, Raleigh, NC 26695-7906, USA (Email: fang@eos.ncsu.edu). This author's work was partially supported by the US Army Research Office Grant No. W911NF-04-D-0003-0002.

‡Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Shatin, NT, Hong Kong (Email: dli@se.cuhk.edu.hk). This author's work was partially supported by Grant CUHK4180/03E, Research Grant Council, Hong Kong.

al. [8] and Boyd and Vandenberghe [6]) can be converted into a convex programming problem by using the log-exp function so that the interior-point algorithms can be developed to solve geometric programs with great efficiency (Kortanek *et al.* [14]). Another example is concerned with the nondifferentiable minimax problem

$$\min_{y\in D}\max_{1\le i\le n} g_i(y),$$

where $g_i(\cdot)$, $i=1,...,n$, are real functions defined on a convex set D in R^m. Since the recession function of the log-exp function is the "max-function"(see Rockafellar [20]), i.e., $\max_{1\le i\le n} x_i = \lim_{\varepsilon\to 0_+}\varepsilon f(\frac{x}{\varepsilon})$ where $f(\cdot)=f_w(\cdot)$ and $w=(1,1,...,1)$, the above nondifferential optimization problem can be approximated by solving the following optimization problem

$$\min_{y\in D}\varepsilon\log\left(\sum_{i=1}^n e^{\frac{g_i(y)}{\varepsilon}}\right).$$

This objective function is differentiable and convex, if every $g_i(y)$ is. Other applications of the log-exp function in optimization can be found in Ben-Tal [1], Ben-Tal and Teboulle [3], Zang [25], Bersekas [4], Polyak [19], Fang [9, 10], Li and Fang [15], Peng and Lin [17], Birbil *et al.* [5], Sun and Li [22, 23, 24]), etc.

It is worth mentioning that the conjugate function of the log-exp function happens to be the well-known Shannon's entropy function ([21]) which plays a vital role in so many fields ranging from the image enhancement to economics and from statistical mechanics to nuclear physics (see, Buck and Macaulay [7] and Fang *et al.* [11]).

We consider in this paper a further generalization of (1.1) in the following form:

$$\Gamma_w(x)=\Psi^{-1}\left(\sum_{i=1}^n w_i\phi_i(x_i)\right) \tag{1.2}$$

where $\phi_i:\Omega\to R$, $i=1,...,n$, are convex, twice differentiable (but not necessarily being strictly increasing) functions defined on an open convex set $\Omega\subset R$, $\Psi:\Omega\to R$ is convex, twice differentiable and strictly increasing, and $w\in R^n_+$ is a given vector. Clearly, $\Upsilon_w(\cdot)$ is a special case of $\Gamma_w(\cdot)$ with $\phi_1=\phi_2=...=\phi_n=\Psi=\phi$. For convenience, in this paper, we still call Γ_w given by (1.2) a generalized mean function, and we call ϕ_i the inner function and Ψ the outer function of Γ_w.

To assure the well-definedness of Γ_w, we naturally require that $\sum_{i=1}^n Cone[\phi_i(\Omega)]\subseteq\Psi(\Omega)$, where $Cone[\phi_i(\Omega)]$ denotes the cone generated by the set $\phi_i(\Omega)$.

As in the case of Υ_w, we would like to derive certain sufficient and necessary conditions for the function Γ_w to be convex. Moreover, we hope to find a systematic way to explicitly construct some classes of convex Γ_w.

It is interesting to point out that Γ_w is by no means a new research subject. In fact, it was essentially studied by W. Fenchel in his lecture notes of "Convex Cones, Sets and Functions" in 1953 [12]. Based on the properties of level sets and characteristic roots of Hessian matrices of functions involved, Fenchel derived some sufficient and necessary conditions for the convexity of the generalized mean function Γ_w. The conditions he derived, however, are rather complicated, and there is no simple test to decide what kind of functions may admit these complicated properties. Unlike Fenchel's approach, our analysis in this paper depends only on the function value, its first derivative, and second derivative to provide a sufficient and necessary condition for Γ_w being convex. The necessary and sufficient condition we derive in this paper can be viewed as a generalization of that in [13] concerning the function (1.1). We can also use related sufficient conditions to explicitly construct concrete examples of convex Γ_w. Moreover, we show how the so-called S^*-regular functions (to be defined in this paper) can be used to construct convex generalized mean functions.

The rest of the paper is organized as follows. In Section 2, we investigate the conditions that assure the convexity of the generalized mean function Γ_w. In Section 3, we identify some classes of functions that satisfy the conditions derived in Section 2, and illustrate how the generalized mean function Γ_w can be explicitly constructed. Conclusions are given in the last section.

2. Necessary and Sufficient Conditions for the Convexity of Γ_w. Let us start with a simple lemma (proof omitted) that shows the inverse of an increasing convex function is concave and increasing.

LEMMA 2.1. *Let Ω be an open convex subset of R and $\Psi : \Omega \to R$ be a real function defined on Ω. Then Ψ is (strictly) convex and strictly increasing if and only if its inverse $\Psi^{-1} : R \to \Omega$ is (strictly) concave and strictly increasing.*

Notice that if $w_i = 0$, for some i, then the term $w_i\phi_i(x)$ can be removed from the expression of $\Gamma_w(x)$, and it suffices to consider Γ_w defined on R^{n-1}. Thus, without loss of generality, we may assume that the vector $w \in R^n_{++}$ throughout the rest of the paper.

To study the convexity of Γ_w, when assuming that ϕ_i, $i = 1, ..., n$, and Ψ^{-1} are twice differentiable, we need to check the properties of its Hessian matrix. Let

$$x_w = \sum_{i=1}^{n} w_i\phi_i(x_i).$$

Since $\frac{\partial x_w}{\partial x_i} = w_i\phi_i'(x_i)$, we have

$$\frac{\partial \Gamma_w}{\partial x_i} = (\Psi^{-1})'(x_w)w_i\phi_i'(x_i).$$

Moreover,

$$\frac{\partial^2 \Gamma_w}{\partial x_i^2} = (\Psi^{-1})''(x_w)(w_i\phi_i'(x_i))^2 + (\Psi^{-1})'(x_w)w_i\phi_i''(x_i),$$

$$\frac{\partial^2 \Gamma_w}{\partial x_i \partial x_j} = (\Psi^{-1})''(x_w)w_iw_j\phi_i'(x_i)\phi_j'(x_j) \quad \text{for} \quad i \neq j.$$

Consequently, the Hessian matrices of Γ_w becomes

$$\frac{\partial^2 \Gamma_w}{\partial x^2} = (\Psi^{-1})'(x_w)\begin{bmatrix} w_1\phi_1''(x_1) & 0 & \dots & 0 \\ 0 & w_2\phi_2''(x_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_n\phi_n''(x_n) \end{bmatrix}$$

$$+ (\Psi^{-1})''(x_w)\begin{bmatrix} w_1\phi_1'(x_1) \\ w_2\phi_2'(x_2) \\ \dots \\ w_n\phi_n'(x_n) \end{bmatrix}[w_1\phi_1'(x_1), w_2\phi_2'(x_2), \dots, w_n\phi_n'(x_n)]. \tag{2.1}$$

Note that when ϕ_i, $i = 1, ..., n$, are all convex and Ψ is convex and increasing, by Lemma 2.1, we see that the first term on the right-hand side of (2.1) is a positive semidefinite matrix multiplied by a positive coefficient $(\Psi^{-1})'(x_w)$, while the second is a rank one matrix multiplied by a negative coefficient $(\Psi^{-1})''(x_w)$.

Some conditions for convexity of the function $\Upsilon_w(x)$ has already been studied in [13] and [2]. We summarize their results here.

THEOREM 2.2. *[13] Under the conditions of $\phi > 0$, $\phi' > 0$ and $\phi'' > 0$, the function $\Upsilon_w(x)$ defined by (1.1) is convex if and only if the following condition holds:*

$$\sum_{i=1}^{n} w_i \frac{[\phi'(x_i)]^2}{\phi''(x_i)} \leq \frac{[\phi'(y)]^2}{\phi''(y)} \quad \textit{for} \quad y = \Upsilon_w(x).$$

Ben-Tal and Teboulle [2] also provided a different sufficient and necessary condition, under certain assumptions, for the convexity of the function $\Upsilon_w(x)$.

THEOREM 2.3. *[2] Let $\phi(t) \in C^3$. $\Upsilon_w(x)$ is convex if and only if $1/\rho(t)$ is convex, where $\rho(t) = -\phi''/\phi'$.*

It is possible to extend the analysis in [2] to derive sufficient conditions for more general situations where the convexity of $\Gamma_w(x)$ defined by (1.2) is considered. Actually, the following result can be proved along the line of the proof of "Lemma 1" and "Theorem 1" in [2].

THEOREM 2.4. *[2] Let $\Psi(t) \in C^3$ and $\phi_i(t) \in C^3$ be strictly increasing and $\rho(t) = -\Psi''(t)/\Psi'(t)$. If $1/\rho(t)$ is convex and $\Psi^{-1}(\phi_i(t))$ is convex, for $i = 1, ..., n$, then $\Gamma_w(x)$ given by (1.2) is convex.*

Note that if ϕ is sufficiently smooth, $1/\rho(t)$ is convex, where $\rho(t) = -\phi''(t)/\phi'(t)$, if and only if its second derivative is nonnegative, i.e.,

$$\left(\frac{1}{\rho}\right)'' = \frac{(\phi'')^3\phi''' - 2\phi'\phi''(\phi''')^2 + \phi'(\phi'')^2\phi''''}{(\phi'')^2} \geq 0.$$

Thus, to check the convexity of $1/\rho(t)$, it is usually needed to check the above inequality involving the third and forth derivative of the function ϕ. Note that Theorem 2.2, however, does not require the third or fourth differentiability of the function ϕ.

In what follows, we generalize the results in Theorem 2.2 for the function $\Gamma_w(x)$. Although the basic idea of our proof is closely related to that of [13], the proof is not straightforward. For completeness, we give a detailed proof for the result.

THEOREM 2.5. *Let $\Omega \subset R$ be open and convex, $\Psi : \Omega \to R$ be convex, twice differentiable and strictly increasing, $\phi_i : \Omega \to R$, $i = 1, ..., n$, be strictly convex and twice differentiable, and $w \in R^n_{++}$ be a given vector. Then the generalized mean function*

$$\Gamma_w(x) = \Psi^{-1}\left(\sum_{i=1}^{n} w_i\phi_i(x_i)\right)$$

is convex on $\Omega^n := \overbrace{\Omega \times \cdots \times \Omega}^{n}$ if and only if

$$\Psi''(y)\left(\sum_{i=1}^{n} w_i \frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}\right) \leq [\Psi'(y)]^2 \quad \textit{for} \quad x \in \Omega^n \textit{ and } y = \Gamma_w(x). \tag{2.2}$$

Moreover, $\Gamma_w(x)$ is strictly convex if and only if the inequality in (2.2) holds strictly.

Proof. Let $y = \Gamma_w(x) = \Psi^{-1}(x_w)$. Then $x_w = \Psi(y)$ and

$$(\Psi^{-1})'(x_w)\Psi'(y) = 1. \tag{2.3}$$

Differentiating both sides with respect to y and use the above relations, we have

$$0 = (\Psi^{-1})''(x_w)[\Psi'(y)]^2 + (\Psi^{-1})'(x_w)\Psi''(y)$$
$$= (\Psi^{-1})''(x_w)[\Psi'(y)]^2 + \frac{\Psi''(y)}{\Psi'(y)}.$$

Therefore,

$$(\Psi^{-1})''(x_w) = -\frac{\Psi''(y)}{[\Psi'(y)]^3}. \tag{2.4}$$

Combining (2.3) and (2.4) yields

$$(\Psi^{-1})'(x_w) + (\Psi^{-1})''(x_w)\sum_{i=1}^{n} w_i \frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)} = \frac{[\Psi'(y)]^2 - \left(\sum_{i=1}^{n} w_i \frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}\right)\Psi''(y)}{[\Psi'(y)]^3}. \tag{2.5}$$

First we prove that $\Gamma_w(x)$ is convex, if (2.2) holds. It suffices to show that the Hessian matrix of $\Gamma_w(x)$ is positive semi-definite.

For any $d \in R^n$ and $x \in \Omega^n$, the Cauchy-Schwartz inequality implies that

$$\left(\sum_{i=1}^{n} w_i\phi_i'(x_i)d_i\right)^2 = \left(\sum_{i=1}^{n}\left[\sqrt{w_i\phi_i''(x_i)}d_i\right]\cdot\sqrt{\frac{w_i}{\phi_i''(x_i)}}\phi_i'(x_i)\right)^2$$
$$\le \left(\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2\right)\left(\sum_{i=1}^{n} w_i\frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}\right).$$

By Lemma 2.1, we know Ψ^{-1} is concave and hence $(\Psi^{-1})''(x_w) \le 0$ for all x_w. Combining this fact with the above inequality, we see that, for any $d \in R^n$,

$$d^T\frac{\partial^2\Gamma_w}{\partial x^2}d$$
$$= (\Psi^{-1})'(x_w)\left(\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2\right) + (\Psi^{-1})''(x_w)\left(\sum_{i=1}^{n} w_i\phi_i'(x_i)d_i\right)^2$$
$$\ge (\Psi^{-1})'(x_w)\left(\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2\right) + (\Psi^{-1})''(x_w)\left(\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2\right)\left(\sum_{i=1}^{n} w_i\frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}\right)$$
$$= \left(\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2\right)\left[(\Psi^{-1})'(x_w) + (\Psi^{-1})''(x_w)\left(\sum_{i=1}^{n}\frac{w_i[\phi_i'(x_i)]^2}{\phi_i''(x_i)}\right)\right]$$
$$= \left(\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2\right)\frac{[\Psi'(y)]^2 - \left(\sum_{i=1}^{n} w_i\frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}\right)\Psi''(y)}{[\Psi'(y)]^3}$$
$$\ge 0.$$

The last equality follows from (2.5) and the last inequality follows from the fact that the first quantity on the right-hand side, i.e., $\sum_{i=1}^{n} w_i\phi_i''(x_i)d_i^2$, is nonnegative, and the second quantity is also nonnegative due to our assumption. Consequently, we have proven that the Hessian matrix $\frac{\partial^2\Gamma_w}{\partial x^2}$ is positive semi-definite, as desired.

Conversely, we would like to show that inequality (2.2) holds, if $\Gamma_w(x)$ is convex. For any vector $0 \neq d \in R^n$, knowing (2.3), (2.4) and the convexity of $\Gamma_w(x)$, we have

$$
\begin{aligned}
0 \leq d^T \frac{\partial^2 \Gamma_w}{\partial x^2} d &= (\Psi^{-1})'(x_w) \left(\sum_{i=1}^n w_i \phi_i''(x_i) d_i^2 \right) + (\Psi^{-1})''(x_w) \left(\sum_{i=1}^n w_i \phi_i'(x_i) d_i \right)^2 \\
&= \frac{1}{\Psi'(y)} \left(\sum_{i=1}^n w_i \phi_i''(x_i) d_i^2 \right) - \frac{\Psi''(y)}{\Psi'(y)^3} \left(\sum_{i=1}^n w_i \phi_i'(x_i) d_i \right)^2 \\
&= \left(\sum_{i=1}^n w_i \phi_i''(x_i) d_i^2 \right) \left[\frac{1}{\Psi'(y)} - \frac{\Psi''(y)}{\Psi'(y)^3} \frac{\left[\sum_{i=1}^n w_i \phi_i'(x_i) d_i\right]^2}{\sum_{i=1}^n w_i \phi_i''(x_i) d_i^2} \right]. \qquad (2.6)
\end{aligned}
$$

Notice that the above inequality holds for any vector $d \in R^n$. In particular, let

$$d_i = \frac{\phi_i'(x_i)}{\phi_i''(x_i) \sum_{k=1}^n w_k \frac{[\phi_k'(x_k)]^2}{\phi_k''(x_k)}}, \; i = 1, ..., n.$$

Then, we have

$$\sum_{i=1}^n w_i \phi_i'(x_i) d_i = 1, \; \sum_{i=1}^n w_i \phi_i''(x_i) d_i^2 = \frac{1}{\sum_{i=1}^n w_i \frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}}.$$

As a result, the inequality (2.6) reduces to

$$0 \leq \left(\frac{1}{\sum_{i=1}^n w_i \frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)}} \right) \left[\frac{1}{\Psi'(y)} - \frac{\Psi''(y)}{\Psi'(y)^3} \left(\sum_{i=1}^n w_i \frac{[\phi_i'(x_i)]^2}{\phi_i''(x_i)} \right) \right]$$

We see that inequality (2.2) indeed holds. The result about strict convexity can be easily checked out. □

Theorem 2.5 generalizes the result of Theorem 2.2 (concerning $\Upsilon_w(x)$) to the more general function $\Gamma_w(x)$, while Theorem 2.4 generalizes the sufficient condition of Theorem 2.3 (concerning $\Upsilon_w(x)$) to the function $\Gamma_w(x)$, under different assumptions. Except for some very simple cases, such as e^t or x^p, these results, however, do not give us a concrete class of functions which can be used to construct specific generalized mean functions. The purpose of the remainder of this paper is to provide a systematic way to identify the desired class of functions. Our analysis here in the paper is based only on the result of Theorem 2.5. We believe that there should also exist some parallel results based on Theorem 2.4.

To this end, two related sufficiency results of Theorem 2.5 are derived below for their convenient usage in constructing convex Γ_w (see next section).

THEOREM 2.6. *Let Ω be an open convex subset of R, $\Psi : \Omega \to R$ be strictly increasing, twice differentiable and convex, $\phi_i : \Omega \to R$, $i = 1, ..., n$, be strictly convex and twice differentiable, and $w \in R^n_{++}$ be a given vector. Assume that there exists a scalar $\alpha \in R$ such that*

$$\alpha \Psi(t) \Psi''(t) \leq [\Psi'(t)]^2 \quad \text{for } t \in \Omega. \qquad (2.7)$$

Then the function Γ_w is convex on Ω^n if

$$\sum_{i=1}^n \frac{w_i [\phi_i'(x_i)]^2}{\phi_i''(x_i)} \leq \alpha \Psi(y) \quad \text{for } x \in \Omega^n, \qquad (2.8)$$

where $y = \Gamma_w(x)$.

Proof. Multiplying both sides of (2.8) by $\Psi''(y)$ and applying (2.7), we see that condition (2.2) holds. The result follows from Theorem 2.5 immediately. ◻

THEOREM 2.7. *Let* Ω *be an open convex subset of* R, $\Psi : \Omega \to R$ *be strictly increasing, twice differentiable and convex,* $\phi_i : \Omega \to R$, $i = 1, ..., n$, *be strictly convex and twice differentiable, and* $w \in R^n_{++}$ *be a given vector. Assume that there exist* $0 \neq \alpha_i \in R$, $i = 1, ..., n$, *holding the same sign such that*

$$\alpha_i \phi_i(t) \phi_i''(t) \geq [\phi_i'(t)]^2 \quad \textit{for } t \in \Omega, \tag{2.9}$$

and there exists an $\alpha \in R$ *such that the inequality (2.7) holds. Then the function* Γ_w *is convex if*

$$\alpha \geq \max_{1 \leq i \leq n} \alpha_i \quad (\textit{ when } \alpha_i > 0 \textit{ for all } i), \tag{2.10}$$

or

$$\alpha \leq \min_{1 \leq i \leq n} \alpha_i \quad (\textit{ when } \alpha_i < 0 \textit{ for all } i). \tag{2.11}$$

Proof. Taking $y = \Gamma_w(x)$, we see two cases.

Case 1: $\alpha_i > 0$ for $i = 1, ..., n$. In this case, (2.9) implies that $\phi_i(t) \geq 0$ for $t \in \Omega$ and (2.10) implies that

$$\sum_{i=1}^n w_i \frac{[\phi_i'(t)]^2}{\phi_i''(x_i)} \leq \sum_{i=1}^n w_i \alpha_i \phi_i(x_i) \leq \left(\max_{1 \leq i \leq n} \alpha_i \right) \sum_{i=1}^n w_i \phi_i(x_i) \leq \alpha \Psi(y).$$

Case 2: $\alpha_i < 0$ for $i = 1, ..., n$. In this case, (2.9) implies that $\phi_i(t) \leq 0$ for $t \in \Omega$ and (2.11) implies that

$$\sum_{i=1}^n w_i \frac{[\phi_i'(t)]^2}{\phi_i''(x_i)} \leq \sum_{i=1}^n w_i \alpha_i \phi_i(x_i) \leq \left(\min_{1 \leq i \leq n} \alpha_i \right) \sum_{i=1}^n w_i \phi_i(x_i) \leq \alpha \Psi(y).$$

Both cases yield (2.8) and the desired result follows from Theorem 2.2. ◻

A special case of $\phi_1(t) = \phi_2(t) = ... = \Psi(t)$ immediately leads to the next result.

COROLLARY 2.8. *Let* Ω *be an open convex set in* R, $\phi : \Omega \to R$ *be a convex, twice differentiable and strictly increasing function, and* $w \in R^n_{++}$ *be a given vector. If there exists an* $\alpha \neq 0$ *such that*

$$[\phi'(t)]^2 = \alpha \phi(t) \phi''(t) \quad \textit{for } t \in \Omega, \tag{2.12}$$

then the function $\Upsilon_w(x) = \phi^{-1}(\sum_{i=1}^n w_i \phi(x_i))$ *is convex on* Ω^n.

This result can also follow directly from the aforementioned Theorem 2.3 (due to Ben-Tal and Teboulle [2]). In fact, it is easy to verify that the relation (2.12) implies that the second derivative of ϕ'/ϕ'' is equal to zero, and thus by Theorem 2.3 the function $\Upsilon_w(x)$ is convex.

Remark 2.1. The functions satisfying a differential inequality such as (2.7) are related to the so-called self-concordant barrier function introduced by Nesterov and Nemirovsky [16]. Recall that a C^3 function $\xi : (0, \infty) \to R$ is said to be self-concordant if ξ is convex and there exists a constant $\mu_1 > 0$ such that

$$|\xi'''(t)| \leq \mu_1 (\xi''(t))^{\frac{3}{2}} \quad \text{for } t \in (0, \infty). \tag{2.13}$$

Moreover, the self-concordant function ξ is called a self-concordant barrier function if there exists a constant $\mu_2 > 0$ such that

$$|\xi'(t)| \leq \mu_2 [f''(t)]^{\frac{1}{2}} \quad \text{for } t \in (0, \infty). \tag{2.14}$$

Combining (2.13) and (2.14) yields

$$\xi'(t)\xi'''(t) \leq \mu[\xi''(t)]^2.$$

This indicates that the first-order derivative function of a self-concordant barrier function, i.e., $g(t) := \xi'(t)$, satisfies the inequality (2.7). A self-concordant function $\xi(\cdot)$ itself may also satisfy an inequality like (2.7) or (2.9).

Remark 2.2. The functions satisfying a differential inequality such as (2.7) also appear in convexity theory. Given a twice differentiable function $\phi(t) > 0$ on its domain Ω, we consider the convexity of the function $h(t) := \frac{1}{\phi(t)}$ on Ω. Notice that

$$h''(t) = \frac{2[\phi'(t)]^2 - \phi(t)\phi''(t)}{[\phi(t)]^3} \quad \text{for} \quad t \in \Omega.$$

Hence the function $h(t) = \frac{1}{\phi(t)}$ is convex if and only if the inequality $\phi(t)\phi''(t) \leq 2[\phi'(t)]^2$ holds on Ω. Moreover, if $\phi(t)\phi''(t) \leq [\phi'(t)]^2$, the convex function $h(t)$ satisfies a reverse inequality, i.e., $h(t)h''(t) \geq [h'(t)]^2$ on Ω.

From this observation, a related question arises. Given a function $\phi(t) > 0$ on Ω and a constant $r > 0$, when will the function $h(t) := \frac{1}{\phi(t)^r}$ become convex and satisfy an inequality such as (2.9)? A straightforward analysis leads to the next result.

THEOREM 2.9. *(i) Let Ω be a convex subset of R and $\phi : \Omega \to (0, \infty)$ be a function. If $\phi(t)\phi''(t) \leq [\phi'(t)]^2$ for $t \in \Omega$, then, for any $r > 0$, the function $h(t) := \frac{1}{\phi(t)^r}$ is convex and $h(t)h''(t) \geq [h'(t)]^2$ for $t \in \Omega$. Conversely, if there exists an $r > 0$ such that $h(t) := \frac{1}{\phi(t)^r}$ is convex and $h(t)h''(t) \geq [h'(t)]^2$ for $t \in \Omega$, then $\phi(t)\phi''(t) \leq [\phi'(t)]^2$ for $t \in \Omega$.*

(ii) Let Ω be a convex subset of R, $\tau > 0$, and $\phi : \Omega \to (\tau, \infty)$ be a function. If $\phi(t)\phi''(t) \leq [\phi'(t)]^2$ for $t \in \Omega$, then, for any scalar $r > 0$ and $T > 0$, the function $h_T(t) := T + \frac{1}{\phi(t)^r}$ is convex and $\alpha h_T(t)h_T''(t) \geq [h_T'(t)]^2$ for $t \in \Omega$, where $\alpha = \frac{1}{T\tau^r+1}$.

Proof. For case (i), it is sufficient to see that

$$h''(t) = \frac{r^2(\phi'(t))^2 + r[(\phi'(t))^2 - \phi(t)\phi''(t)]}{\phi(t)^{r+2}},$$

and

$$h(t)h''(t) - [h'(t)]^2 = \frac{r[(\phi'(t))^2 - \phi(t)\phi''(t)]}{\phi(t)^{2(r+1)}}.$$

For case (ii), it is easy to verify that $h_T''(t) = h''(t)$ and

$$\left(\frac{1}{T\phi(t)^r + 1}\right) h_T(t)h_T''(t) - [h_T'(t)]^2 = \frac{r[(\phi'(t))^2 - \phi(t)\phi''(t)]}{\phi(t)^{2(r+1)}}.$$

Then the desired result follows. □

The above results indicate that if we have a function ϕ satisfying the inequality (2.7) with $\alpha = 1$, then we may construct a function h from ϕ such that h satisfies the converse differentiable inequality $\alpha h(t)h''(t) \geq [h'(t)]^2$ for some constant α. Moreover, if we take a T-translation of the value of the function h, then the resulting function satisfies the converse differentiable inequality with an α that can be reduced to be smaller than any threshold given in (0,1) provided a suitable choice of $T > 0$. This fact will be used near the end of Section 3.

3. Constructing convex generalized mean functions Γ_w. In this section, we develop procedures to identify some classes of functions that satisfy inequality (2.7) and/or inequality (2.9) so that we have building blocks for constructing concrete convex function $\Gamma_w(x)$. First, we give a result that identifies functions satisfying the equation (2.12). Obviously, this class of functions satisfies both inequalities (2.7) and (2.9).

THEOREM 3.1. *Let Ω be an open set in R and $\phi : \Omega \to R$ be a convex, twice differentiable and strictly increasing function satisfying equation (2.12) with a constant $\alpha \neq 0$. Then,*

(i) when $\alpha = 1$, ϕ is in the form of $\phi(t) = \gamma e^{\frac{t}{\beta}}$ for some $\gamma > 0$ and $\beta > 0$.

(ii) when $0 < \alpha \neq 1$ with $v^* := \sup_{t\in\Omega} \frac{1-\alpha}{\alpha} t$ being finite, ϕ is in the form of

$$\phi(t) = \gamma \left(\frac{\alpha - 1}{\alpha} t + \beta \right)^{\frac{\alpha}{\alpha-1}}$$

for some $\gamma > 0$ and $\beta \geq v^*$.

(iii) when $\alpha < 0$ with $u^* := \sup_{t\in\Omega} \frac{\alpha-1}{\alpha} t$ being finite, ϕ is in the form of

$$\phi(t) = -\gamma \left(\beta - \frac{\alpha - 1}{\alpha} t \right)^{\frac{\alpha}{\alpha-1}}$$

for some $\gamma > 0$ and $\beta \geq u^*$.

Note that results (i) and (ii) were pointed out in [2] and [13] and result (iii) can be easily derived. The above result leads to the following consequence related to Υ_w.

COROLLARY 3.2. *The following functions can be used to explicitly construct a convex generalized mean function $\Upsilon_w(x) = \phi^{-1}(\sum_{i=1}^n w_i \phi(x_i))$ over Ω^n:*

(i) $\phi(t) = \gamma e^{\frac{t}{\beta}}$ over $\Omega = R$ with $\gamma > 0$ and $\beta > 0$.

(ii) $\phi(t) = \gamma \left(\frac{1}{p} t + \beta\right)^p$ over $\Omega = (\eta, \infty)$ with $p > 1$, $\gamma > 0$ and $\beta \geq -\frac{\eta}{p}$.

(iii) $\phi(t) = \frac{\gamma}{(\beta - \frac{1}{p} t)^p}$ over $\Omega = (-\infty, \eta)$ with $p > 0$, $\gamma > 0$ and $\beta \geq -\frac{\eta}{p}$.

(iv) $\phi(t) = -\gamma(\beta - \frac{1}{p} t)^p$ over $\Omega = (-\infty, \eta)$ with $0 < p < 1$, $\gamma > 0$ and $\beta \geq \frac{\eta}{p}$.

Again, results (i) and (ii) were given in [2] and [13] and results (iii) and (iv) can be easily derived. The functions listed in Corollary 3.2 actually form a complete basis in the sense that the function ϕ in case (i) satisfies condition (2.12) with $\alpha = 1$; the function ϕ in case (ii) satisfies condition (2.12) with $\alpha = \frac{p}{p-1} > 1$; the function ϕ in case (iii) satisfies condition (2.12) with $\alpha = \frac{p}{p+1} \in (0, 1)$; and the function ϕ in (iv) satisfies condition (2.12) with $\alpha = \frac{p}{p-1} < 0$.

We now try to identify some class of functions that satisfy inequalities (2.7) and/or (2.9). For simplicity, we only consider convex, twice differentiable, strictly increasing functions ϑ on $\Omega = (0, \infty)$. Let us first define the following four categories of such functions:

$$\mathcal{U}_1 = \{\vartheta : \text{ There exists } \alpha \in R \text{ such that } \alpha\vartheta(t)\vartheta''(t) \geq [\vartheta'(t)]^2 \text{ for } t \in \Omega\};$$
$$\mathcal{U}_2 = \{\vartheta : \text{ There exists } \alpha \in R \text{ such that } \alpha\vartheta(t)\vartheta''(t) \leq [\vartheta'(t)]^2 \text{ for } t \in \Omega\};$$

$\mathcal{U}_3 = \{\vartheta :$ There exist $\alpha_1 \le \alpha_2$ such that $\alpha_1 \vartheta(t)\vartheta''(t) \le [\vartheta'(t)]^2 \le \alpha_2 \vartheta(t)\vartheta''(t)$ for $t \in \Omega\}$;

$\mathcal{U}_4 = \{\vartheta :$ There exists $\alpha \in R$ such that $\alpha\vartheta(t)\vartheta''(t) = [\vartheta'(t)]^2$ for all $t \in \Omega\}$.

It is evident that

$$\mathcal{U}_4 \subset \mathcal{U}_3 \subset (\mathcal{U}_2 \cap \mathcal{U}_1).$$

As pointed out in Theorem 3.1, the class $\mathcal{U}_4$ can be given explicitly. By allowing $\alpha_1 \neq \alpha_2$, we show that $\mathcal{U}_3$ is much broader than $\mathcal{U}_4$. In fact, many convex functions with certain regularities fall into the category $\mathcal{U}_3$. To start, we introduce a new class of functions with certain regularity properties.

Definition 3.3. *A convex, twice differentiable, strictly increasing function* $\delta(t) : (0,\infty) \to R$ *is called an* S^**-regular function if (i)* $\delta(t)$ *vanishes at* $t = 0$ *in the sense of*

$$\lim_{t\to 0_+} \delta(0) = \lim_{t\to 0_+} \delta'(0) = \lim_{t\to 0_+} \delta''(0) = 0;$$

and (ii) there exist positive constants $0 < \beta_1 \le \beta_2$, $p \ge 1$ *and* $q \ge 1$ *such that*

$$\beta_1[(t+1)^{p-1} - (t+1)^{-1-q}] \le \delta''(t) \le \beta_2[(t+1)^{p-1} - (t+1)^{-1-q}],\ t > 0. \tag{3.1}$$

Note that condition (3.1) actually implies the strict convexity of an S^*-regular function on $(0,\infty)$. In particular, setting $\beta_1 = \beta_2$, condition (3.1) reduces to an equation

$$\delta''(t) = (t+1)^{p-1} - (t+1)^{-1-q}. \tag{3.2}$$

Taking integration twice and noting that $\lim_{t\to 0_+} \delta(0) = \lim_{t\to 0_+} \delta'(0) = 0$, the unique solution to equation (3.2) is

$$\Delta_{p,q}(t) = \frac{(t+1)^{p+1} - 1}{p(p+1)} - \frac{(t+1)^{1-q} - 1}{q(q-1)} - \frac{p+q}{pq}t \quad \text{for } p \ge 1 \text{ and } \ q > 1. \tag{3.3}$$

In addition, since $\lim_{q\to 1_+}[1 - (t+1)^{1-q}]/(q-1) = \ln(t+1)$, we have

$$\Delta_{p,1}(t) = \frac{(t+1)^{p+1} - 1}{p(p+1)} + \ln(t+1) - \frac{p+1}{p}t \quad \text{for } p \ge 1. \tag{3.4}$$

Taking $p = 1$ in (3.4), we have

$$\Delta_{1,1}(t) = \frac{(t+1)^2 - 1}{2} + \ln(t+1) - 2t = \frac{1}{2}t^2 - t + \ln(t+1). \tag{3.5}$$

Moreover, taking $p = 1$ and $q = 2$ in (3.3) yields

$$\Delta_{1,2}(t) = \frac{1}{2}\left[(t+1)^2 - (t+1)^{-1} - 3t\right]. \tag{3.6}$$

In terms of this particular solution $\Delta_{p,q}(t)$, condition (3.1) can be written as

$$\beta_1 \Delta''_{p,q}(t) \le \delta''(t) \le \beta_2 \Delta''_{p,q}(t). \tag{3.7}$$

By integrating and noting that $\lim_{t\to 0_+} \delta'(0) = \lim_{t\to 0_+} \delta(0) = 0$, we further have

$$\beta_1 \Delta'_{p,q}(t) \le \delta'(t) \le \beta_2 \Delta'_{p,q}(t) \tag{3.8}$$

and

$$\beta_1 \Delta_{p,q}(t) \le \delta(t) \le \beta_2 \Delta_{p,q}(t). \tag{3.9}$$

Therefore, we can see that the class of S^*-regular functions is quite broad. Later, by using (3.7)-(3.9), we show that S^*-regular functions fall into the category $\mathcal{U}_3$.

It is worth mentioning that for any $p \ge 1, q > 1$ (including the case of $q \to 1_+$) the S^*-regular function $\Delta_{p,q}(t)$ is not self-concordant. In fact, the function $\Delta_{p,q}(t)$ does not satisfy the inequality (2.13) since $\delta''(t) \to 0$ and $\delta'''(t) \to p+q$ as $t \to 0_+$.

S^*-regular functions are somewhat analogous to (but different from) the self-regular functions defined in [18]. As we have mentioned above, S^*-regular functions are not self-concordant. The class of self-regular functions, however, has a large overlap with self-concordant functions. In what follows, we display the relation among the first and second derivatives of S^*-regular functions, which shows that any S^*-regular function belongs to the category $\mathcal{U}_3$. It should be mentioned that the relations among the first and second derivatives for self-regular function have been studied in [18].

THEOREM 3.4. *Let $\delta(t) : (0, \infty) \to R$ be S^*-regular on $(0, \infty)$. Then there exist $c_2 \ge c_1 > 0$ such that*

$$c_1 \le \frac{\delta(t)\delta''(t)}{[\delta'(t)]^2} \le c_2 \text{ for all } t \in (0, \infty), \tag{3.10}$$

i.e., the function $\delta(t) \in \mathcal{U}_3$.

Proof. We show that an S^*-regular function $\Delta_{p,q}(t)$ satisfies the property (3.10). Actually, we have

$$\frac{\Delta_{p,q}(t)\Delta''_{p,q}(t)}{[\Delta'_{p,q}(t)]^2} = \frac{\left(\frac{(t+1)^{p+1}-1}{p(p+1)} - \frac{(t+1)^{1-q}-1}{q(q-1)} - \frac{p+q}{pq}t\right)[(t+1)^{p-1} - (t+1)^{-1-q}]}{\left(\frac{(t+1)^p}{p} + \frac{(t+1)^{-q}}{q} - \frac{p+q}{pq}\right)^2}.$$

Dividing the numerator and denominator of the right-hand side of the above equation by $(t+1)^{2p} = (t+1)^{p+1}(t+1)^{p-1}$, we have

$$\frac{\Delta_{p,q}(t)\Delta''_{p,q}(t)}{[\Delta'_{p,q}(t)]^2} = \frac{\left(\frac{1-(t+1)^{-(p+1)}}{p(p+1)} + \frac{(t+1)^{-(p+1)}-(t+1)^{-(p+q)}}{q(q-1)} - \frac{(p+q)t}{pq(t+1)^{(p+1)}}\right)\left(1 - \frac{1}{(t+1)^{(p+q)}}\right)}{\left(\frac{1}{p} + \frac{1}{q(t+1)^{(p+q)}} - \frac{p+q}{pq(t+1)^p}\right)^2}.$$

Therefore,

$$\lim_{t\to\infty} \frac{\Delta_{p,q}(t)\Delta''_{p,q}(t)}{[\Delta'_{p,q}(t)]^2} = \frac{p}{p+1}. \tag{3.11}$$

Since $\Delta''_{p,q}(t) = (t+1)^{p-1} - (t+1)^{-1-q}$, we have $\lim_{t\to 0_+} \Delta'''_{p,q}(t) = p+q$. Since $\Delta''_{p,q}(t) \to 0$, $\Delta'_{p,q}(t) \to 0$ and $\Delta_{p,q}(t) \to 0$ as $t \to 0_+$, we have

$$\lim_{t\to 0_+} \frac{(\Delta''_{p,q}(t))^2}{\Delta'_{p,q}(t)} = \lim_{t\to 0_+} \frac{[(\Delta''_{p,q}(t))^2]'}{[\Delta'_{p,q}(t)]'} = \lim_{t\to 0_+} \frac{2\Delta''_{p,q}(t)\Delta'''_{p,q}(t)}{\Delta''_{p,q}(t)} = 2(p+q).$$

Hence, we have

$$\lim_{t\to 0_+} \frac{\Delta_{p,q}(t)}{2\Delta'_{p,q}(t)\Delta''_{p,q}(t)} = \lim_{t\to 0_+} \frac{\Delta'_{p,q}(t)}{2[\Delta''_{p,q}(t)]^2 + 2\Delta'_{p,q}(t)\Delta'''_{p,q}(t)} = \frac{1}{6(p+q)}.$$

Using the above relations, we further have

$$
\begin{aligned}
\lim_{t\to 0_+}\frac{\Delta_{p,q}(t)\Delta''_{p,q}(t)}{[\Delta'_{p,q}(t)]^2} &= \lim_{t\to 0_+}\frac{\Delta'_{p,q}(t)\Delta''_{p,q}(t)+\Delta_{p,q}(t)\Delta'''_{p,q}(t)}{2\Delta'_{p,q}(t)\Delta''_{p,q}(t)}\\
&= \frac{1}{2}+\lim_{t\to 0_+}\frac{\Delta_{p,q}(t)}{2\Delta'_{p,q}(t)\Delta''_{p,q}(t)}\lim_{t\to 0_+}\Delta'''_{p,q}(t)=\frac{2}{3}. \qquad (3.12)
\end{aligned}
$$

Notice that $\Delta_{p,q}(t)>0, \Delta''_{p,q}(t)>0$ and $\Delta'_{p,q}(t)>0$ in $(0,\infty)$. From (3.11) and (3.12), we can see by continuity that there exist two constants $\mu_2 \geq \mu_1 > 0$ such that

$$\mu_1 \leq \frac{\Delta_{p,q}(t)\Delta''_{p,q}(t)}{[\Delta'_{p,q}(t)]^2} \leq \mu_2 \quad \text{for } t\in(0,\infty).$$

Together with (3.7) through (3.9), this implies that an S^*-regular function $\delta(t)$ satisfies the following inequality:

$$0<\mu_1\beta_1 \leq \frac{\delta(t)\delta''(t)}{[\delta'(t)]^2} \leq \beta_2\mu_2,$$

Therefore, (3.10) holds with $c_1 := \mu_1\beta_1$ and $c_2 := \mu_2\beta_2$. □

A fact that should be pointed out here is that new functions in $\mathcal{U}_1$ or $\mathcal{U}_2$ can be constructed by using the basic operations (addition, multiplication, division and composition) on known functions. The proof of the following fact is omitted.

LEMMA 3.5. *(i) If $\phi:(0,\infty)\to(0,\infty)$, $\phi\in\mathcal{U}_1$ with $\alpha=\alpha_1$ and $\varphi:(0,\infty)\to(0,\infty)$, $\varphi\in\mathcal{U}_1$ with $\alpha=\alpha_2$, then $\phi+\varphi\in\mathcal{U}_1$ with $\alpha=2\max\{\alpha_1,\alpha_2\}$.*

(ii) If $\phi:(0,\infty)\to(0,\infty)$, $\phi\in\mathcal{U}_1$ with $\alpha_1\in(0,1]$ and $\varphi:(0,\infty)\to(0,\infty)$, $\varphi\in\mathcal{U}_1$ with $\alpha_2\in(0,1]$, then the multiplicative function $\phi(t)\cdot\varphi(t)\in\mathcal{U}_1$ with $\alpha=1$. Similarly, if $\phi\in\mathcal{U}_2$ with $\alpha_1\geq 1$ and $\varphi\in\mathcal{U}_2$ with $\alpha_2\geq 1$, then $\phi(t)\cdot\varphi(t)\in\mathcal{U}_2$ with $\alpha=1$.

(iii) If $\phi:(0,\infty)\to(0,\infty)$, $\phi\in\mathcal{U}_2$ with $\alpha_1\geq 1$ and $\varphi:(0,\infty)\to(0,\infty)$, $\varphi\in\mathcal{U}_1$ with $\alpha_2\in(0,1]$, then the function $\frac{\phi}{\varphi}\in\mathcal{U}_2$ with $\alpha=1$. Similarly, if $\phi\in\mathcal{U}_1$ with $\alpha_1\in(0,1]$ and $\varphi\in\mathcal{U}_2$ with $\alpha_2\geq 1$, then $\frac{\phi}{\varphi}\in\mathcal{U}_1$ with $\alpha=1$.

(iv) Let $\varphi:(0,\infty)\to\Omega_1\subset R$ and $\phi:\Omega_1\to(0,\infty)$ be two convex functions. If $\phi\in\mathcal{U}_1$ with $\alpha>0$, then the composite function $(\phi\circ\varphi)(t)=\phi(\varphi(t))\in\mathcal{U}_1$ with the same constant α.

The next result shows that the composite functions of e^t belong to $\mathcal{U}_3$.

LEMMA 3.6. *Denote the exponential function e^t by $\exp(t)$ and the composition of m ($m\geq 1$) exponential functions by*

$$\theta_m(t) := (\overbrace{\exp\circ\exp\circ\cdots\circ\exp}^{m})(t).$$

Then

$$\frac{1}{m}\theta_m(t)\theta''_m(t) \leq [\theta'_m(t)]^2 \leq \theta_m(t)\theta''_m(t) \quad \text{for } t\in R. \qquad (3.13)$$

Proof. Let $\alpha_m(t) := [\theta'_m(t)]^2/(\theta_m(t)\theta''_m(t))$ for $t\in R$. Since $\alpha_1(t)\equiv 1$, we can prove the right-hand side inequality of (3.13) using (iv) of Lemma 3.5 and mathematical induction. For the left-hand side inequality, notice that

$$\theta'_m(t)=\theta_m(t)\theta'_{m-1}(t),\ \ \theta''_m(t)=\theta_m(t)(\theta'_{m-1}(t))^2+\theta_m(t)\theta''_{m-1}(t) \quad \text{for } t\in R.$$

This indicates that

$$\alpha_m(t) = \frac{1}{1 + \frac{1}{\alpha_{m-1}(t)\theta_{m-1}(t)}} > \frac{1}{1 + \frac{1}{\alpha_{m-1}(t)}} \quad \text{for } t \in R.$$

It is easy to check that $\alpha_2(t) \in (\frac{1}{2}, 1)$. The desired result follows by induction. □

To construct examples of the convex function Γ_w, Theorem 2.7 tells us that it suffices to find functions satisfying the inequalities (2.7) and (2.9) and compare their α values. The next result is to estimate the α values, or equivalently, to estimate the values of c_1 and c_2 in (3.10). For simplicity, we use the S^*-regular functions with $p = 1$ and $q = 1, 2$ to estimate required c_1 and c_2. In fact, we have the following result. Its proof is omitted here.

LEMMA 3.7. *The S^*-regular functions $\Delta_{1,1}(t)$ and $\Delta_{1,2}(t)$ given by (3.5) and (3.6), respectively, satisfy condition (3.10) with $c_1 = \frac{1}{2}$ and $c_2 = \frac{2}{3}$, that is,*

$$\frac{3}{2}\Delta_{1,1}(t)\Delta''_{1,1}(t) \le [\Delta'_{1,1}(t)]^2 \le 2\Delta_{1,1}(t)\Delta''_{1,1}(t), \tag{3.14}$$

$$\frac{3}{2}\Delta_{1,2}(t)\Delta''_{1,2}(t) \le [\Delta'_{1,2}(t)]^2 \le 2\Delta_{1,2}(t)\Delta''_{1,2}(t) \tag{3.15}$$

for $t \in (0, \infty)$.

We now give the last result on how to construct some convex functions Γ_w.

THEOREM 3.8. *Let Ω be an open convex subset of R.*

(i) Let $\phi : \Omega \to (0, \infty)$ be a convex, twice differentiable, strictly increasing function on Ω. If $\phi(t)\phi''(t) \le [\phi'(t)]^2$ for $t \in \Omega$, then the generalized mean function

$$\Gamma_w^{(1)}(x) := \phi^{-1}\left(\sum_{i=1}^{n} \frac{w_i}{\phi(x_i)^r}\right)$$

is convex on Ω^n for any given $w \in R^n_{++}$ and $r > 0$.

(ii) Let $\kappa > 0$ be a constant and $\phi : \Omega \to (\kappa, \infty)$ be a convex, twice differentiable, strictly increasing function satisfying the inequality $\phi(t)\phi''(t) \le [\phi'(t)]^2$ for $t \in \Omega$. Then, for any given $w \in R^n_{++}$ and $T > 0, r > 0$, the function

$$\Gamma_w^{(2)}(x) := \overbrace{\ln \circ \ln \circ \cdots \circ \ln}^{\ell}\left(\sum_{i=1}^{n} w_i \left(T + \frac{1}{\phi(x_i)^r}\right)\right)$$

is convex on Ω^n for any positive integer $\ell \le T\kappa^r + 1$.

In fact, result (i) comes from part (i) of Theorem 2.9 and Theorem 2.7. Result (ii) follows from Lemma 3.6 and Theorem 2.7, and part (ii) of Theorem 2.9. In fact, it suffices to take the inner function $h_T(t) = T + \frac{1}{\phi(t)^r}$ and outer function $\theta_m(t)$, as defined in Lemma 3.6, whose inverse function is given by $\overbrace{\ln \circ \ln \cdots \circ \ln}^{m}(t)$.

The above result partially answers the following interesting question: *Given a convex function, how many times of* log*-transformations can be applied while retaining the convexity?*

Using Theorems 2.7, 2.9, 3.8 and Lemma 3.7, we have the following examples of convex Γ_w.

Example 3.1.

(i) $\Delta_{1,j}^{-1}\left[\sum_{i=1}^{n}\frac{1}{\Delta_{1,j}(x_i)^r}\right]$,

(ii) $\ln\left(\sum_{i=1}^{n}\frac{1}{\Delta_{1,j}(x_i)^r}\right)$,

(iii) $\overbrace{\ln\circ\ln\circ\cdots\circ\ln}^{\ell\le m+1}\left(\sum_{i=1}^{n}\left(m+e^{-rx_i}\right)\right),\ x\in(0,\infty)^n.$

(iv) $\overbrace{\ln\circ\ln\circ\cdots\circ\ln}^{\ell}\left[\sum_{i=1}^{n}\left(m+\frac{1}{\Delta_{1,1}(x_i)^r}\right)\right],\ x\in(\tau,\infty)^n, \ell\le m\Delta_{1,1}(\tau)^r+1, \tau>0.$

It follows from Corollary 3.2 that the function x^p over $(0,\infty)$ satisfies (2.12) with $\alpha=\frac{p}{p-1}$. Hence, when $1<p\le 2$, we have $\alpha\ge 2$, and when $1<p\le\frac{29}{17}$, we have $\alpha\ge\frac{29}{12}\ge\frac{9}{4}$. By Theorems 3.7, both $\Delta_{1,2}(t)$ and $\Delta_{1,1}(t)$ satisfy condition (2.9) with $\alpha=2$. From Theorem 2.7, we see the functions below are examples of convex Γ_w.

Example 3.2. Let $1<p\le 2$ and $\delta_i(t)=\Delta_{1,2}(t)$ or $\Delta_{1,1}(t)$, for $t\in(0,\infty)$ and $i=1,...,n$. Then $\Gamma_w(x)=(\sum_{i=1}^{n}w_i\delta_i(x_i))^{\frac{1}{p}}$ is convex on $(0,\infty)^n$.

Before closing this section, we briefly illustrate a possible application of involving function Γ_w in the regularization method for solving a nonlinear programming problem:

$$\min\{f_0(x): x\in C\}.$$

For simplicity, we assume that C is a convex set and f_0 is a convex function. Let $\mu>0$ be a positive parameter. Given a strictly convex function Γ_w, we consider the following problem:

$$\min\{f_0(x)+\mu\Gamma_w(x): x\in C\}.$$

This problem becomes a strictly convex programming problem with a unique solution, denoted by $x(\mu)$, which comprises of a continuation trajectory $\{x(\mu):\mu>0\}$. Under suitable conditions of f_0,Ψ and ϕ, this trajectory becomes bounded. In this case, by setting $\mu\to 0$, any accumulation point of $x(\mu)$, as $\mu\to 0$, is a solution to the original problem. Thus, a path-following algorithm can be designed to follow this trajectory to achieve the solution of the original problem. The performance of such path following algorithm certainly depends on the choice of the function Γ_w with regularity conditions.

4. Conclusions. In this paper, we have further extended the theoretical foundation for the generalized mean function. We have established a necessary and sufficient condition for such a generalization to be convex. Moreover, a systematic way to explicitly construct convex Γ_w has been developed. To this end, the concept of S^*-regular functions has been introduced. It should be noted that any S^*-regular function is not self-concordant [16].

Acknowledgments. We would like to thank the two anonymous referees for their insightful comments which help improve significantly the results and presentation of the paper. Especially, we are grateful to one of the referees who brought the references [2] and [13] to our attention, and provided “Theorem 2.4” in this paper.

REFERENCES

[1] A. Ben-Tal, *The entropic penalty approach to stochastic programming,* Mathematics of Operations Research, 10 (1985), pp. 263-279.

[2] A. Ben-Tal and M. Teboulle, *Expected utility, penalty functions, and duality in stochastic nonlinear programming,* Management Science, 32 (1986), pp. 1445-1466.

[3] A. Ben-Tal and M. Teboulle, *A smoothing technique for nondifferentiable optimization problems,* in Optimization, Lecture Note in Mathematics 1405, Springer Verlag, pp. 1-11, 1989.

[4] D.P. Bertsekas, *Constrained Optimization and Lagrangian Multiplier Methods,* Academic Press, New York, 1982.

[5] S.I. Birbil, S.-C. Fang, J. Frenk and S. Zhang, *Recursive approximate of the high dimensional MAX function,* Operations Research Letters, 33 (2005), pp.450-458.

[6] S. Boyd and L. Vandenberghe, *Introdunction to Convex Optimization with Engineering Applications,* Stanford University, Stanford, CA, 1997.

[7] B. Buck and V.A. Macaulay, *Maximum Entropy in Action: A Collection of Expository Essays*, Oxford University Press, London UK, 1991.

[8] R.J. Duffin, E.L. Peterson and C. Zener, *Geometric Programming - Theory and Applications,* Wiley, 1967.

[9] S.-C. Fang, *An unconstrained convex programming view of linear programming,* Mathematical Methods of Operations Research, 36 (1992), pp. 149-161.

[10] S.-C. Fang and H. S. J. Tsao, *On the entropic perturbation and exponential penalty methods for linear programming,* Journal of Optimization Theory and Applications, 89 (1996), pp. 461-466.

[11] S.-C. Fang, J.R. Rajasekera and H. Tsao, *Entropy Optimization and Mathematical Programming,* Kluwer Academic Publishers, Boston, MA, 1997.

[12] W. Fenchel, *Convex Sets, Cones, and Functions,* Lectures in Preceton University, Princeton, New Jersey, 1953.

[13] G. Hardy, J.L. Littlewood and G. Pólya, *Inequalities,* Cambridge University Press, 1934.

[14] K.O. Kortanek, X. Xu and Y. Ye, *An infeasible interior-point algorithm for solving primal and dual geometric programs,* Mathematical Programming, 76 (1996), pp. 155-181.

[15] X.-S. Li and S.-C. Fang, *On the regularization method for solving min-max problems with applications,* Mathematical Methods of Operations Research, 46 (1997), pp. 119-130.

[16] Y. Nesterov and A. Nemirovsky, *Interior-point Polynomial Methods in Convex Programming,* SIAM Studies 13, Philadelphia, PA, 1994.

[17] J. Peng and Z. Lin, *A non-interior continuation method for generalized linear complementarity problems,* Mathematical Programming, 86 (1999), pp. 533-563.

[18] J. Peng, C. Roos, and T. Terlaky, *Self-Regularity: A new paradigm for primal-dual interior-point algorithms,* Princeton University Press, 2002.

[19] R. A. Polyak, *Smooth optimization method for minimax problems,* SIAM J. Control and Optimization, 26 (1988), pp. 1274-1286.

[20] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[21] C.E. Shannon, *A Mathematical Theory of Communication,* Bell System Technical Journal, 27 (1948), pp. 379–423 and 623–656.

[22] X. L. Sun and D. Li, *Value-estimation function method for constrained global optimization,* Journal of Optimization Theory and Applications, 102 (1999), pp. 385-409.

[23] X. L. Sun and D. Li, *Logarithmic-exponential penalty formulation for integer programming,* Applied Mathematics Letters, 12(1999), pp. 73-77.

[24] X. L. Sun and D. Li, *Asymptotic strong duality for bounded integer programming: A logarithmic-exponential dual formulation,"* Mathematics of Operations Research, 25 (2000), pp. 625-644.

[25] I. Zang, *A smoothing technique for min-max optimization,* Mathematical Programming, 19 (1980), pp. 61-77.