Convex quadratic programming with weighted entropic perturbation

by
Vasile Preda and Costel Bălcău

Abstract
We extend the unconstrained convex programming approach to solving the standard-form quadratic programming problem with weighted entropic perturbation. We construct a geometric dual problem for the proposed problem and we prove their weak and strong duality theorems.

Key Words: Convex quadratic programming, entropic perturbation, geometric programming method, duality theorems.

2000 Mathematics Subject Classification: Primary 90C20, Secondary 90C25, 94A17.

1 Introduction
The method of minimizing the weighted logarithm deviation from a reference measure subject to some given moments was used by Guiaşu [10] to classify known probability distributions. For a guide to further references and applications see Guiaşu [9, 11, 12].

In this paper we study the quadratic programming problem with weighted entropic perturbation

\[(P(\mu)) : \min f_\mu(x) = \frac{1}{2}x^\top Dx + c^\top x + \sum_{j=1}^{n} \mu_j x_j \ln x_j \quad \text{s.t.} \]

\[Ax = b; \quad x \geq 0,\]

where \(x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n\) is the variable vector for the problem, \(n \in \mathbb{N}^*\), \(c = (c_1, \ldots, c_n)^\top \in \mathbb{R}^n\) and \(b = (b_1, \ldots, b_m)^\top \in \mathbb{R}^m\) are known vectors, \(A = (a_{ij})_{i \in \{1, \ldots, m\}} \in \mathbb{R}^{m \times n}\) is a known matrix, \(m \in \mathbb{N}^*, m \leq n, D \in \mathbb{R}^{n \times n}\) is a known

*Dedicated to Professor Silviu Guiaşu on his 70th birthday.*
diagonal matrix having the diagonal elements $d_1, \ldots, d_n \geq 0$, $\mu_1, \ldots, \mu_n > 0$ are perturbation parameters (weights), and $\mu = (\mu_1, \ldots, \mu_n)^T \in \mathbb{R}^n$. We use the well-known convention $0 \ln 0 = 0$.

We remark that in the particular case when $\mu_1 = \mu_2 = \ldots = \mu_n$ we regain the standard quadratic programming with entropic perturbation [8].

In [5, 6, 7, 8, 14] Fang, Rajasekera and Tsao developed the entropic perturbation method for solving linear and quadratic programming problems using Erlander’s geometric entropic programming technique [4].

Using this method, in Section 2 we construct a geometric dual problem of problem $(P(\mu))$, and in Section 3 we describe their duality. As special cases we recover duality from quadratic programming with entropic perturbation [8], from linear programming with weighted entropic perturbation [1], and from maximizing the weighted entropy subject to some given constraints of moments type [10].

We mention that some interesting extensions of the entropic perturbation method for linear or quadratic programming can be also found in [2, 3, 13, 16].

2 The construction of the dual

Because of the entropic perturbation, the problem $(P(\mu))$ seems more complicated than the standard-form quadratic programming problem. However, this perturbation allows us to construct an unconstrained dual convex programming problem.

We will assume that the problem $(P(\mu))$ has at least one interior feasible solution $x > 0$. This assumption is called the interior point assumption.

Remark 2.1. Under the interior point assumption the problem $(P(\mu))$ is consistent. More than that, as

$$
\lim_{x_j \to 0} x_j \ln x_j = 0, \quad \lim_{x_j \to -\infty} \left( \frac{1}{2} d_j x_j^2 + c_j x_j + \mu_j x_j \ln x_j \right) = +\infty
$$

and the function $x_j \ln x_j$ is strictly decreasing on the interval $[0, \frac{1}{e}]$ and strictly convex for each $j \in \{1, \ldots, n\}$, it follows that the problem $(P(\mu))$ has a finite optimum and a unique optimal solution $x^*(\mu) > 0$.

Let $j \in \{1, \ldots, n\}$ be arbitrary and fixed. For each positive function $\varphi_j(y)$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ and for each $x_j > 0$, applying the logarithmic inequality

$$
\ln \frac{\varphi_j(y)}{x_j} \leq \frac{\varphi_j(y)}{x_j} - 1,
$$

with equality if and only if

$$
\varphi_j(y) = x_j,
$$

(1)
Convex quadratic programming

it follows that
\[ \mu_j \ln \frac{\varphi_j(y)}{x_j} \leq \frac{d_j \left[ x_j - \varphi_j(y) \right]^2}{2x_j} + \frac{\mu_j \varphi_j(y)}{x_j} - \mu_j, \quad (2) \]

with equality if and only if the relation (1) holds.

The inequality (2) can be rewritten as
\[ \frac{d_j x_j}{2} + \mu_j \ln x_j \geq \left[ d_j \varphi_j(y) + \mu_j \ln \varphi_j(y) \right] - \frac{d_j \varphi_j^2(y)}{2x_j} - \frac{\mu_j \varphi_j(y)}{x_j} + \mu_j. \quad (3) \]

As the function
\[ F_j : (0, +\infty) \to \mathbb{R}, \quad F_j(t) = d_j t + \mu_j \ln t \]
is bijective, the following equation
\[ d_j \varphi_j(y) + \mu_j \ln \varphi_j(y) = \sum_{i=1}^{m} a_{ij} y_i - c_j - \mu_j \quad (4) \]
has a unique positive solution \( \varphi_j(y) \), for every \( y \in \mathbb{R}^m \) and \( j \in \{1, \ldots, n\} \).

Now the inequality (3) becomes
\[ \frac{d_j x_j}{2} + c_j + \mu_j \ln x_j \geq \sum_{i=1}^{m} a_{ij} y_i - \frac{d_j \varphi_j^2(y)}{2x_j} - \mu_j \varphi_j(y). \]

Multiplying both members by \( x_j > 0 \) and summing after \( j \in \{1, \ldots, n\} \) we get
\[ f_\mu(x) \geq \sum_{i=1}^{m} y_i \sum_{j=1}^{n} a_{ij} x_{j} - \frac{1}{2} \sum_{j=1}^{n} d_j \varphi_j^2(y) - \sum_{j=1}^{n} \mu_j \varphi_j(y). \quad (5) \]

Passing to the limit \( x_j \searrow 0 \) it follows that this inequality is true for every \( x_j \geq 0, \) \( j \in \{1, \ldots, n\} \).

If the vector \( x = (x_1, \ldots, x_n)^T \) satisfies the restriction \( Ax = b \), the inequality (5) can be rewritten as
\[ f_\mu(x) \geq \sum_{i=1}^{m} b_i y_i - \frac{1}{2} \sum_{j=1}^{n} d_j \varphi_j^2(y) - \sum_{j=1}^{n} \mu_j \varphi_j(y). \quad (6) \]

In this way we can define the geometric dual problem of problem \((P(\mu))\), namely
\[ (D(\mu)) : \max_{y \in \mathbb{R}^m} g_\mu(y) = b^T y - \frac{1}{2} \sum_{j=1}^{n} d_j \varphi_j^2(y) - \sum_{j=1}^{n} \mu_j \varphi_j(y). \]
3 Duality results

In this section we will establish duality theorems between problems \((P(\mu))\) and \((D(\mu))\). The next theorem follows directly from (6).

**Theorem 3.1** (Weak duality). If \(x\) and \(y\) are feasible solutions of problems \((P(\mu))\) and \((D(\mu))\) respectively, then \(f_\mu(x) \geq g_\mu(y)\).

Because one has equality in (6) if and only if \(x_j = \varphi_j(y)\), \(\forall j \in \{1, \ldots, n\}\), we obtain the following result.

**Theorem 3.2.** Let \(y^* = (y_1^*, \ldots, y_m^*) \in \mathbb{R}^m\) and \(x^* = (x_1^*, \ldots, x_n^*) \in \mathbb{R}^n, x^* > 0\) be such that the following relation holds

\[
d_j x_j^* + \mu_j \ln x_j^* = \sum_{i=1}^{m} a_{ij} y_i^* - c_j - \mu_j, \quad \forall j \in \{1, \ldots, n\}.
\]

If \(x^*\) satisfies the restriction \(Ax^* = b\), then \(x^*\) is the optimal solution of problem \((P(\mu))\) and \(y^*\) is an optimal solution of problem \((D(\mu))\). Moreover, in this case

\[
f_\mu(x^*) = g_\mu(y^*),
\]

i.e. problems \((P(\mu))\) and \((D(\mu))\) have the same optimal value.

To prove the strong duality theorem we will need the following result.

**Lemma 3.1.** The function \(g_\mu(y)\) is concave.

**Proof:** We have \(g_\mu(y) = \Theta(y) + \sum_{j=1}^{n} \Psi_j(y)\), where \(\Theta(y) = \sum_{i=1}^{m} b_i y_i, \forall y \in \mathbb{R}^m\), and

\[
\Psi_j(y) = -\frac{1}{2} d_j \varphi_j^2(y) - \mu_j \varphi_j(y), \quad \forall y \in \mathbb{R}^m, \quad \forall j \in \{1, \ldots, n\}.
\]

Obviously the function \(\Theta(y)\) is linear. Hence it is enough to prove that the function \(\Psi_j(y)\) is concave for every \(j \in \{1, \ldots, n\}\). We have

\[
\frac{\partial \Psi_j}{\partial y_k}(y) = -(d_j \varphi_j(y) + \mu_j) \frac{\partial \varphi_j}{\partial y_k}(y), \quad \forall k \in \{1, \ldots, m\}.
\] (7)

For each \(j \in \{1, \ldots, n\}\), define the function \(\phi_j : (0, +\infty) \times \mathbb{R}^m \to \mathbb{R}\) by

\[
\phi_j(x_j, y) = d_j x_j + \mu_j \ln x_j - \sum_{i=1}^{m} a_{ij} y_i + c_j + \mu_j.
\]

It follows from (4) that \(\phi_j(\varphi_j(y), y) = 0, \forall y \in \mathbb{R}^m\). From

\[
\frac{\partial \phi_j}{\partial y_k}(\varphi_j(y), y) = 0, \quad \forall k \in \{1, \ldots, m\}, \quad \forall y \in \mathbb{R}^m
\]
we derive that
\[ \frac{\partial \varphi_j}{\partial y_k}(y) = \frac{a_{kj} \varphi_j(y)}{d_j \varphi_j(y) + \mu_j}, \quad \forall k \in \{1, \ldots, m\}. \quad (8) \]

Then, from (7) we deduce that the Hessian of \( \Psi_j(y) \) has the form
\[ H_{\Psi_j}(y) = r_j(y) A^{(j)} A^{(j)^\top}, \]
where
\[ r_j(y) = -\frac{\varphi_j(y)}{d_j \varphi_j(y) + \mu_j}, \quad A^{(j)} = \left( a^{(j)_{il}} \right)_{i \in \{1, \ldots, m\}, \ l \in \{1, \ldots, n\}}, \]
\[ a^{(j)_{il}} = \begin{cases} a_{ij}, & \text{if } l = j \\ 0, & \text{if } l \neq j \end{cases}. \]
Obviously, \( r_j(y) < 0 \), and then the Hessian \( H_{\Psi_j}(y) \) is negative semi-definite, which means that \( \Psi_j(y) \) is concave.

**Remark 3.1.** It follows from the proof of Lemma 3.1 that if \( \text{rank} A = m \), then \( g_\mu(y) \) is strictly concave, which guarantees the uniqueness of the optimal solution of the dual problem \((D(\mu))\).

**Theorem 3.3 (Strong duality).** Assume that the primal problem \((P(\mu))\) satisfies the interior point assumption, \( \text{rank} A = m \), and the dual problem \((D(\mu))\) has feasible solutions. Then the two problems have the same optimal value. Moreover, if \( y^*(\mu) \) is the optimal solution of problem \((D(\mu))\), then the vector \( x^*(\mu) \) given by the equation
\[ d_j x^*_j(\mu) + \mu_j \ln x^*_j(\mu) = \sum_{i=1}^{m} a_{ij} y^*_i(\mu) - c_j - \mu_j, \quad \forall j \in \{1, \ldots, n\} \quad (9) \]
is the optimal solution of problem \((P(\mu))\) and \( f_\mu(x^*(\mu)) = g_\mu(y^*(\mu)) \).

**Proof:** The primal problem \((P(\mu))\) can be written in the form
\[ \inf_{x \in \mathbb{R}^n} \left\{ \tilde{f}(x) - h(Ax) \right\}, \]
where
\[ \tilde{f}(x) = \begin{cases} f_\mu(x), & \text{if } x \geq 0 \\ +\infty, & \text{otherwise} \end{cases}, \quad \forall x \in \mathbb{R}^n, \]
and
\[ h(w) = -\delta(w|\{b\}) = \begin{cases} 0, & \text{if } w = b \\ -\infty, & \text{if } w \neq b \end{cases}, \quad \forall w \in \mathbb{R}^m, \]
\( \delta(\cdot|\{b\}) \) being the indicator function of the set \( \{b\} \). Obviously, \( \tilde{f} \) is a closed proper convex function and \( h \) is a closed proper concave function. Using Fenchel Duality Theorem ([15], p. 332) we deduce that the Fenchel dual of the problem \((P(\mu))\) is
\[ \sup_{y \in \mathbb{R}^m} \left\{ h^*(y) - \tilde{f}^*(A^\top y) \right\}, \quad (10) \]
where the functions $h^*$ and $\tilde{f}^*$ are the Fenchel conjugates of $h$, and $\tilde{f}$ respectively. According to the definition of the Fenchel conjugate, we have

$$\tilde{f}^*(u) = \sup_{x \geq 0} \left\{ x^\top u - f(x) \right\}$$

$$= \sum_{j=1}^n \left[ -\frac{1}{2} d_j(x_j^*)^2 + (u_j - c_j)x_j^* - \mu_j x_j^* \ln x_j^* \right].$$

where $x_j^* > 0$ is the unique solution of the equation

$$d_j x_j + \mu_j \ln x_j = u_j - c_j - \mu_j. \quad (11)$$

It follows that $\tilde{f}^*(u) = \sum_{j=1}^n \left[ \frac{1}{2} d_j x_j^* \right].$

For $u = A^\top y$, according to (4) the equation (11) has a unique solution $x_j^* = \varphi_j(y)$. Hence

$$\tilde{f}^*(A^\top y) = \sum_{j=1}^n \left[ \frac{1}{2} d_j \varphi_j^2(y) + \mu_j \varphi_j(y) \right], \forall y \in \mathbb{R}^m. \quad (12)$$

Obviously,

$$h^*(y) = -\delta^*(-y|\{b\}) = b^\top y, \forall y \in \mathbb{R}^m. \quad (13)$$

Using (10), (12) and (13) we deduce that the Fenchel dual of the problem $(P(\mu))$ is even the geometric dual problem $(D(\mu))$. It follows from our hypothesis and the Fenchel Duality Theorem that the problems $(P(\mu))$ and $(D(\mu))$ have optimal solutions and the same optimal value.

Let now $y^*(\mu)$ be the optimal solution of the problem $(D(\mu))$ and consider the vector $x^*(\mu)$ given by the equation (9). According to (9) and (4) it follows that $x_j^*(\mu) = \varphi_j(y^*(\mu)) > 0, \forall j \in \{1, \ldots, n\}$.

Using the optimality of the dual solution $y^*(\mu)$ and the relation (8) from the proof of Lemma 3.1 we have

$$0 = \frac{\partial g_\mu}{\partial y_k}(y^*(\mu)) = b_k - \sum_{j=1}^n a_{kj} \varphi_j(y^*(\mu)) = b_k - \sum_{j=1}^n a_{kj} x_j^*(\mu), \forall k \in \{1, \ldots, m\}.$$

Then $Ax^*(\mu) = b$. By Theorem 3.2, we conclude that $x^*(\mu)$ is the optimal solution of problem $(P(\mu))$ and $f_\mu(x^*(\mu)) = g_\mu(y^*(\mu))$. \hfill \square

**Remark 3.2.** For $\mu_1 = \mu_2 = \ldots = \mu_n = \mu$ one recovers the duality from the standard quadratic programming with entropic perturbation [8].
Remark 3.3. In the absence of the quadratic term, that is for \( D = 0 \), one can solve the equation (4) and gets explicitly the function

\[
\varphi_j(y) = \exp \left[ \frac{1}{\mu_j} \left( \sum_{i=1}^{m} a_{ij} y_i - c_j \right) - 1 \right], \quad \forall j \in \{1, \ldots, n\}.
\]

In this way one recovers the duality from the linear programming with weighted entropic perturbation [1].

 Remark 3.4. If \( c = 0 \), \( D = 0 \) and \( \mu \) is a finite probability distribution, one recovers the duality from the maximizing of weighted entropy subject to some given constraints of moments type [10].

References


Received: 28.07.2008.

University of Bucureşti,
Faculty of Mathematics and Informatics
E-mail: preda@fmi.unibuc.ro

University of Piteşti,
Faculty of Mathematics and Informatics
E-mail: cbalcau01@linux.math.upit.ro