An FPTAS for Minimizing a Class of Low-Rank Quasi-Concave Functions over a Convex Set

Vineet Goyal* R. Ravi†

Abstract

We consider minimizing a class of low rank quasi-concave functions over a convex set and give a fully polynomial time approximation scheme (FPTAS) for the problem. The algorithm is based on a binary search for the optimal objective value which is guided by solving a polynomial number of linear minimization problems over the convex set with appropriate objective functions. Our algorithm gives a \((1 + \epsilon)\)-approximate solution that is an extreme point of the convex set and therefore, has direct applications to combinatorial 0-1 problems for which the convex hull of feasible solutions is known, such as shortest paths, spanning trees and matchings in undirected graphs. Our results also extend to maximization of low-rank quasi-convex functions over a convex set.

Key words: Quasi-concave programming, non-linear programming, non-convex programming, polynomial approximation schemes.

*Operations Research Center, Massachusetts Institute of Technology, Cambridge MA 02139. Email: goyalv@mit.edu. Part of the work was done as a graduate student at Carnegie Mellon University and was partially supported by NSF grant CCF-0728841.

†Tepper School of Business, Carnegie Mellon University, Pittsburgh PA 15213. Email: ravi@andrew.cmu.edu. Supported in part by NSF grant CCF-0728841. Corresponding Author.
1 Introduction

Concave minimization over a convex domain has been an extensively studied optimization problem (see [8, 17]) and finds important applications that include problems with economies of scale and the fixed charge problem [7, 16]. In addition, several important optimization problems can be formulated as a concave minimization problem such as 0-1 integer programming [19], quadratic assignment [14, 3], bilinear programming [10], linear complementary problem [15]. Therefore, the concave minimization problem includes a broad range of problems and is thus of considerable interest.

Concave minimization is NP-hard even for the special case of minimizing a concave quadratic function over a hypercube [6]. While several solution approaches including enumeration [2, 4] and cutting plane methods [21] (see [8] for a complete list of references) have been studied, the concave minimization problem still remains a very hard problem computationally.

Since the general concave minimization problem is hard, special cases of the problem have been of interest. Konno et al. [11] consider the special case of low-rank concave quadratic minimization problem and propose a tabu-search heuristic to solve the problem. However, the heuristic does not guarantee a global optimum and there is no worst-case bound on the objective value achieved by the heuristic. Porembski [18] considers a generalization of the low-rank concave quadratic functions where the objective function is non-linear for a small number of variables and propose a cutting plane solution approach. However, to the best of our knowledge there is no work that gives an efficient algorithm with a provable worst-case bound on the objective value of the solution, even for the special case of minimizing low-rank concave functions.

In this paper, we consider the problem of minimizing a low-rank quasi-concave function over a compact convex set.

Definition 1.1 A function $f : D \rightarrow \mathbb{R}$ over a convex set $D$ is concave if for all $x, y \in D$ and $\alpha \in [0, 1],$

$$f(\alpha x + (1 - \alpha)y) \geq \alpha f(x) + (1 - \alpha)f(y).$$

$f$ is quasi-concave if and only if the upper level sets $U_\lambda = \{x \in D | f(x) \geq \lambda\}$ for all $\lambda \in \mathbb{R}$ are convex.

Note that a concave function $f$ is also quasi-concave; however, the converse is not true. For example, $f(x_1, x_2) = x_1 x_2$ is a quasi-concave function in $\mathbb{R}^2_+$ but is not concave [1]. However, note that $f$ is not quasi-concave on $\mathbb{R}^2$.

Therefore, our problem is a generalization of the problem considered in Konno et al. [11] and Porembski [18]. Following [12, 18], we define the rank of a quasi-concave function as:

Definition 1.2 The rank of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the smallest number $k$ such that for all $x \in \mathbb{R}^n, f(x) = g(a_1^T x, a_2^T x, \ldots, a_k^T x)$ for some function $g : \mathbb{R}^k \rightarrow \mathbb{R}$ and linearly independent vectors $a_1, a_2, \ldots, a_k \in \mathbb{R}^n$.

Remark: If $f$ is quasi-concave on a convex set $D$, then $g$ is quasi-concave on $D_k = \{(a_1^T x, \ldots, a_k^T x) | x \in D\}$.

We consider the problem: $\min \{f(x) | x \in D\}$ where $f$ is a quasi-concave function $f$ of rank $k$ and $D \subset \mathbb{R}^n_+$ is a compact convex set. Since $f$ has rank $k$, there exists a function $g : \mathbb{R}^k \rightarrow \mathbb{R}$ such that $f(x) = g(a_1^T x, \ldots, a_k^T x)$ for some $a_1, \ldots, a_k \in \mathbb{R}^n$. Therefore, we can reformulate our problem as:

$$\min \{g(y) | \exists x \in D \text{ such that } y_i = a_i^T x, \forall i = 1, \ldots, k\}.$$
Since $\mathcal{D}$ is a convex set, $\mathcal{D}'$ is also convex.

In wake of the above reformulation, we can assume w.l.o.g. that the quasi-concave function $f$ is of the form $f : \mathbb{R}_+^k \rightarrow \mathbb{R}$ and the minimization problem ($\Pi(f, \mathcal{D})$) can be stated as

$$\min \{ f(x) | (x, y) \in \mathcal{D} \}$$

We assume that $f$ satisfies the following conditions:

(P1) The gradient $\nabla f \geq 0$ for all $x \in \mathcal{D}$, i.e. partial derivative $f_{x_i} = \frac{\partial f}{\partial x_i} \geq 0$ for all $i = 1, \ldots, k$.

(P2) $f(\alpha x) \leq \alpha^c f(x)$ for all $\alpha \geq 1$ and some constant $c$.

Conditions (P1) and (P2) are satisfied by a large class of quasi-concave functions that include linear multiplicative functions [13] and utility functions from microeconomics theory such as Leontief utilities and Cobb-Douglas utilities [5].

2 \hspace{1em} (1 + \epsilon)-Approximation Algorithm

We present a fully polynomial time approximation scheme (FPTAS) for the problem $\Pi(f, \mathcal{D})$ in this section. Let $l_i = \min_{(x,y) \in \mathcal{D}} x_i$ and $u_i = \max_{(x,y) \in \mathcal{D}} x_i$ and let $(x_i^{(i)}, y_i^{(i)}), (x_u^{(i)}, y_u^{(i)}) \in \mathcal{D}$ be the minimizers for $l_i$ and $u_i$ respectively for all $i = 1, \ldots, k - 1$. We assume that the diameter of the set $\mathcal{D}$ is at most $R$ and $l_i > 0$ for all $i \in [k - 1]$. We construct a polynomial size grid for the subspace $[l_1, u_1] \times \cdots \times [l_{k-1}, u_{k-1}]$. For a given $\epsilon > 0$, let $n_i = \lceil \log(1+\epsilon) \frac{u_i}{l_i} \rceil$ and $G_i = \{ l_i \cdot (1 + \epsilon)^j | j = 0, 1, \ldots, n_i \}$. We consider the grid of points given by $G = G_1 \times G_2 \times \cdots \times G_{k-1}$.

The basic algorithm can be described as follows. Let $\bar{z}$ and $\underline{z}$ denote the current upper bound and lower bound respectively on $f$. We do a binary search for the optimal objective value $z^*$ between $\underline{z}$ and $\bar{z}$. For a particular value $\lambda$ of the objective function value, let $\mathcal{P}_\lambda = \{ x = (x_1, \ldots, x_k) | (x_1, \ldots, x_{k-1}) \in G \text{ and } f(x) = \lambda \}$. We solve a set of linear minimization programs over the convex set $\mathcal{D}$ and based on the optimal objective values of these linear programs, we are able to distinguish whether $z^* < \lambda$ or $\lambda \leq z^*(1 + \epsilon)$. A formal description of the algorithm is given in Figure 1. We initialize $\underline{z}$ and $\bar{z}$ as follows.

$$\underline{z} \leftarrow f(l_1, l_2, \ldots, l_k) \hspace{1em} (1)$$

$$\bar{z} \leftarrow \min_{i=1}^{k} \{ f(x_i^{(i)}), f(x_u^{(i)}) \} \hspace{1em} (2)$$

Since $\nabla f \geq 0$, $\underline{z}$ is a valid lower bound of the objective value.

To prove that the algorithm is correct, we need to show that for each possible value $\lambda$ of the objective function, we either update the upper bound $\bar{z}$ in step (2c) or the lower bound $\underline{z}$ in step (3). It is clear that $\bar{z}$ is updated only if we find a feasible solution $(x_v^*, y_v^*) \in \mathcal{D}$ with objective value better than the current upper bound $\bar{z}$. The lower bound $\underline{z}$ update in step (3) is not obvious and we prove that it is valid in the following lemma.

Lemma 2.1 For any $\lambda \in \mathbb{R}$, if for all $v \in \mathcal{P}_\lambda$, $z_v^* \geq (\nabla f(v))^T v$, then $\lambda \leq z^*(1 + \epsilon)^c$ where $z^*$ is the optimal value of $\Pi(f, \mathcal{D})$. 

2
Input: a rank-\(k\) quasi-concave function \(f\) satisfying (P1) and (P2), a compact convex set \(D \subset \mathbb{R}^n_+\) and \(\epsilon > 0\).

while \((z > (1 + \epsilon)z)\)

1. \(\lambda \leftarrow \frac{z^+ + z^-}{2}\) and let \(P_\lambda\) be the set of grid points as defined.

2. For each \(v \in P_\lambda\),
   
   (a) Consider the linear function \((\nabla f(v))^Tx\)
   
   (b) Solve the linear minimization problem \(LP(v)\): \(z^*_v = \min \{ (\nabla f(v))^Tx \mid (x, y) \in D \}\) and let the minimum be attained at the extreme point \((x^*_v, y^*_v) \in D\)
   
   (c) If \(z > f(x^*_v)\), \(z \leftarrow f(x^*_v)\) and \((x_A, y_A) \leftarrow (x^*_v, y^*_v)\).

3. If for all \(v \in P_\lambda\), \(z^*_v \geq (\nabla f(v))^Tv\), then \(z \leftarrow \lambda (1 + \epsilon)^T\).

Return \((x_A, y_A)\).

Figure 1: Algorithm \(A\) for Minimizing a low-rank Quasi-concave Function

Proof: Suppose \((x^*, y^*) \in D\) is the optimal solution that minimizes \(f(x)\). There exists \(\beta \in G_{k-1}\) such that \(\frac{\beta_i}{1 + \epsilon} \leq x^*_i < \beta_i\) for all \(i = 1, \ldots, k - 1\). Let \(\beta_k\) be such that \(v = (\beta_1, \ldots, \beta_{k-1}, \beta_k) \in P_\lambda\). Consider the linear minimization problem \(LP(v)\).

\[
z^*_v = \min \{ (\nabla f(v))^Tx \mid (x, y) \in D \}\]

We know that \(z^*_v \geq (\nabla f(v))^Tv\). Therefore,

\[
\sum_{i=1}^{k} f_{x_i}(v) \beta_i \leq z^*_v \tag{3}
\]

\[
\leq \sum_{i=1}^{k} f_{x_i}(v) x^*_i \tag{4}
\]

\[
\leq \sum_{i=1}^{k-1} f_{x_i}(v) \beta_i + f_{x_k}(v) x^*_k \tag{5}
\]

Inequality (4) follows from the optimality of \(z^*_v\) and inequality (5) follows from the fact the \(f_{x_i}(v) \geq 0, x^*_i \leq \beta_i, \forall i = 1, \ldots, k - 1\). Therefore, \(\beta_k \leq x^*_k\). Now,

\[
\lambda = f(v = (\beta_1, \ldots, \beta_k)) \leq f((1 + \epsilon)x^*) \leq (1 + \epsilon)^e f(x^*)
\]

Here, the last inequality follows from the fact that \(f(\alpha x) \leq \alpha^e f(x), \forall \alpha \geq 1\) and some constant \(e\).

We now prove the correctness of the algorithm by showing that we either update the upper bound or the lower bound in each iteration of the algorithm.
Lemma 2.2 At the end of each iteration with \( \lambda = \frac{z + z^*}{\lambda} \) in the algorithm, either the upper bound \( z \) is updated to less than \( \lambda \) or the lower bound \( z^* \) is updated to \( \lambda(1 + \epsilon)^c \) in step (3).

Proof: If \( z^*_v \geq (\nabla f(v))^T v \) for all \( v \in \mathcal{P}_\lambda \), then we know that the lower bound \( z^* \) is updated to \( \lambda(1 + \epsilon)^c \) using Lemma 2.1. Suppose this is not the case. Then, there exists a \( v \in \mathcal{P}_\lambda \) such that \( z^*_v < (\nabla f(v))^T v \). In this case we show that we can update the upper bound \( z \).

Consider the upper level set \( U_\lambda = \{ x \in \mathbb{R}_+^k | f(x) \geq \lambda \} \). Since \( f \) is quasi-concave, \( U_\lambda \) is convex. Therefore, \( (\nabla f(v))^T x \geq (\nabla f(v))^T v \) is a valid supporting hyperplane for \( U_\lambda \) [20].

Now, \( z^*_v = (\nabla f(v))^T x^*_v < (\nabla f(v))^T v \) which implies that \( x^*_v \notin U_\lambda \) and \( f(x^*_v) < \lambda \). Therefore, the upper bound \( z \) is updated to \( f(x^*_v) \) which is less than \( \lambda \).

The following lemma bounds the running time of the algorithm.

Lemma 2.3 The running time of the algorithm \( A \) is polynomial in input size of the problem and \( \frac{1}{\epsilon} \).

Proof: We do a binary search for the objective value between the initial upper and lower bounds (\( z \) and \( z^* \)) until \( z \leq z(1 + \epsilon) \). If \( z^0(z^0) \) denote the initial upper (lower) bound on the objective value, then we would consider \( O(\log |z^0|) \) different possibilities for the objective value in the algorithm. For each possible value (say \( \lambda \)), we solve \( O(|\mathcal{P}_\lambda|) \) linear minimization problems over the convex set \( \mathcal{D} \). Let \( R = \max_{i=1}^k \frac{u_i}{l_i} \); note that \( R \) is polynomial in the input size of the problem. Therefore, the size of the discretization for coordinate \( i \), \( |G_i| = O((\log(1+\epsilon)R) \approx O(\frac{\log R}{\epsilon}) \). Thus, the size of the grid \( G \) and the set of points \( \mathcal{P}_\lambda \) is \( O \left( \frac{\log R}{\epsilon} \right)^{k-1} \). This shows that our algorithm \( A \) solves

\[
O \left( \log \left| \frac{z^0}{\epsilon z^0} \right| \cdot \frac{(\log R)^{k-1}}{\epsilon^{k-1}} \right) \]

linear optimization problems which is polynomial in the input size and \( \frac{1}{\epsilon} \).
Thus, we have the following theorem.

**Theorem 2.4** There is a fully polynomial time approximation scheme (FPTAS) for the problem \( \Pi(f,D) : \min \{f(x)| (x,y) \in D \} \) where \( D \) is a convex set and \( f : \mathbb{R}^k \rightarrow \mathbb{R} \) is a quasi-concave function of rank-\( k \) for some constant \( k \) such that \( f_{x_i} \geq 0 \) for all \( i = 1, \ldots, k \) and \( f(\alpha x) \leq \alpha^c f(x) \) for all \( \alpha \geq 1 \) and some constant \( c \).

Note that \( f \) is minimized at an extreme point of \( D \) since for all \( u,v \in D \) and \( 0 \leq \alpha \leq 1 \),

\[
f(\alpha u + (1-\alpha)v) \geq \min \{f(u), f(v)\}.
\]

Furthermore, our algorithm \( A \) outputs a solution that is an extreme point of \( D \).

### 2.1 Applications to 0-1 Combinatorial Problems

We show applications of our algorithm to combinatorial 0-1 problems. For the case when the convex set \( D \) is a polytope, we obtain a basic optimal solution for the linear minimization problems in each iteration and maintain that the best feasible solution is a vertex of \( D \) throughout the algorithm. Therefore, the algorithm \( A \) gives an approximate solution that is a vertex of \( D \) and thus, has direct applications to several combinatorial 0-1 problems. Consider a combinatorial problem

\[
\min \{c^T x | x \in S \} \tag{6}
\]

where \( S \subset \{0,1\}^n \) is the set of feasible solutions of the combinatorial problem and \( c \in \mathbb{R}^n_+ \). Our FPTAS can be adapted to approximately solve the problem: \( \min \{f(x)| x \in S \} \) where \( f \) is a low rank quasi-concave function that satisfies conditions (P1) and (P2) if we are able to solve (6) efficiently. Therefore, our result applies when \( S \) is the set of \( s-t \) paths, spanning trees, perfect matchings or \( s-t \) cuts in an undirected graph.

**Example.** Consider the following linear multiplicative function: \( f(x) = \Pi_{i=1}^k (a_i^T x + b_i) \) where \( b_i \geq 0, \forall i \in [k] \). It is known that \( f \) is quasi-concave \([13]\). Let \( g : \mathbb{R}^k \rightarrow \mathbb{R} \) where \( g(x) = \Pi_{i=1}^k (x_i + b_i) \). Therefore, for any \( x \in \mathbb{R}^n \), \( f(x) = g(a_1^T x, \ldots, a_k^T x) \) and hence \( f \) has rank \( k \). Furthermore, the function \( g \) satisfies property (P1) as \( g_{x_i} = \Pi_{j \neq i} (x_j + b_j) \geq 0, \forall j = 1, \ldots, k \) and property (P2) as \( g(\alpha x) \leq \alpha^k g(x) \) for all \( \alpha \geq 1 \). Therefore, we can give an FPTAS for the problem: \( \min \{f(x)| x \in S \} \) for combinatorial problems where we are able to solve (6) efficiently. This generalizes the result of Kern and Woeginger \([9]\) who consider minimizing the product of two linear functions over a polytope and give an FPTAS for the same.

### 3 Quasi-convex Maximization

We show that our results extend to the problem of maximizing a low rank quasi-convex function.

**Definition 3.1** A function \( f : D \rightarrow \mathbb{R} \) over a convex set \( D \) is convex if for all \( x,y \in D \) and \( \alpha \in [0,1] \),

\[
f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)
\]

\( f \) is quasi-convex if and only if the lower level sets \( L_\lambda = \{ x \in D | f(x) \leq \lambda \} \) for all \( \lambda \in \mathbb{R} \) are convex.
We consider the following maximization problem: \( \max \{ f(x) | (x, y) \in D \} \) where \( D \) is a compact convex set and \( f: \mathbb{R}^k_+ \to \mathbb{R} \) is a quasi-convex function that satisfies (i) (P1) \( f_{x_i} \geq 0 \) for all \( i = 1, \ldots, k \) and (P2) \( f(\alpha x) \leq \alpha^c f(x) \) for all \( \alpha \geq 1 \) and some constant \( c \).

If a function \( f \) is quasi-convex then \(-f\) is clearly quasi-concave. Therefore, in general the problem of maximizing a quasi-convex function \( f \) over a convex set \( D \) is equivalent to minimizing the quasi-concave function \(-f\). However, the algorithm presented in Section 2 gives an FPTAS for a class of low rank quasi-concave functions that satisfy the conditions (P1) and (P2). If \( f \) satisfies (P1) and (P2), then \(-f\) does not satisfy these properties and thus, we can not directly use Theorem 2.4 to obtain an FPTAS for the quasi-convex maximization problem.

Using the notation from Section 2, let \( l_i = \min_{(x,y) \in D} x_i \) and \( u_i = \max_{(x,y) \in D} x_i \) and let \((x_{l(i)}^{(i)}, y_{l(i)}^{(i)})\) and \((x_{u(i)}^{(i)}, y_{u(i)}^{(i)})\) be the minimizers for \( l_i \) and \( u_i \) respectively for all \( i = 1, \ldots, k - 1 \). We assume that \( l_i > 0 \) for all \( i \in [k - 1] \). We construct a polynomial size grid \( G \) for the subspace \([l_1, u_1] \times \ldots \times [l_{k-1}, u_{k-1}]\) as before. Also, for any possible objective value \( \lambda \), let \( \mathcal{P}_\lambda = \{x = (x_1, \ldots, x_k) | (x_1, \ldots, x_{k-1}) \in G \text{ and } f(x) = \lambda\} \). We initialize the lower bound \( z \) and the upper bound \( \overline{z} \) as follows.

\[
\begin{align*}
    z & \leftarrow \max_{i=1}^{k} \{ f(x_{l(i)}^{(i)}), f(x_{u(i)}^{(i)}) \} \\
    \overline{z} & \leftarrow f(u_1, u_2, \ldots, u_k)
\end{align*}
\]  

The algorithm is formally described in Figure 3.

**Input:** a rank-\( k \) quasi-convex function \( f \) satisfying (P1) and (P2), a compact convex set \( D \subset \mathbb{R}^n_+ \) and \( \epsilon > 0 \).

while \( (\overline{z} - z) > (1 + \epsilon)z \)

1. \( \lambda \leftarrow \frac{\overline{z} + z}{2} \) and let \( \mathcal{P}_\lambda \) be the set of grid points as defined.

2. For each \( v \in \mathcal{P}_\lambda \),

   (a) Consider the linear function \((\nabla f(v))^T x\).

   (b) Solve the linear maximization problem \(LP(v): z_{v}^* = \max \{ (\nabla f(v))^T x | (x, y) \in D \}\) and let the maximum be attained at the extreme point \((x_v^*, y_v^*) \in D\).

   (c) If \( z < f(x_v^*), \overline{z} \leftarrow f(x_v^*) \) and \((x_A, y_A) \leftarrow (x_v^*, y_v^*)\).

3. If for all \( v \in \mathcal{P}_\lambda, z_v^* \leq (\nabla f(v))^T v \), then \( \overline{z} \leftarrow \lambda(1 + \epsilon)z \).

Return \((x_A, y_A)\).

Figure 3: Algorithm \( A_{max} \) for Maximizing a low-rank Quasi-convex Function

It is clear that the lower bound update in step (2c) is correct. Using arguments similar to Lemma 2.1 and Lemma 2.2, we can show that the upper bound update in step (3) is valid and in each iteration of
the algorithm, we either update the lower bound or the upper bound. Therefore, we obtain the following theorem.

**Theorem 3.2** There is a fully polynomial time approximation scheme (FPTAS) for the problem \( \Pi(f, D) : \max\{f(x)|(x, y) \in D\} \) where \( D \) is a compact convex set and \( f : \mathbb{R}^k_+ \to \mathbb{R} \) is a quasi-convex function of rank-\( k \) for some constant \( k \) that satisfies (i) (P1) \( f_{x_i} \geq 0 \) for all \( i = 1, \ldots, k \) and (ii) (P2) \( f(\alpha x) \leq \alpha^c f(x) \) for all \( \alpha \geq 1 \) and some constant \( c \).

4 Conclusions

In this paper, we consider minimizing a low-rank quasi-concave function \( f \) that satisfies certain properties and obtain an FPTAS for the problem. Since \( f \) is minimized at an extreme point of the convex set, there exists a linear objective function that is minimized at the same extreme point as \( f \). Our algorithm can be interpreted as an efficient search of such a linear objective function where we search over the gradient of \( f \) at a polynomial number of points. This search procedure is very general and it would be interesting to see if the idea can be extended to other non-convex optimization problems.

References


