RR-73-21

A Generalized Programming Solution to a Convex Programming Problem with a Homogeneous Objective*

George B. Dantzig

At a recent IIASA Seminar, Youri A. Rozanov [1] posed the following convex programming problem:

Find Min $F(x)$ such that:

$$F(x) = cx + \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2} \tag{1}$$

$$Ax = b, \ x \geq 0 \quad , \tag{2}$$

where $0 \leq k \leq n$, $x = (x_1,\ldots,x_n)$, $c = (c_1,\ldots,c_n)$, and A is an $m \times n$ matrix of rank m. Our objective is to provide an efficient algorithm for solving such a problem.

Initial B

We begin by constructing a non-singular $m \times m$ matrix

$$B = [P_1, P_2, \ldots, P_m] \quad , \tag{3}$$

where each P_i is generated by some $x = x^i$ by the relation

$$P_i = Ax^i \quad , \quad x^i \geq 0 \quad , \qquad \text{for } i = (1,\ldots,m) \quad . \tag{4}$$

It is not required that each x^i satisfy $Ax^i = b$, but we do require that if we solve

$$B\lambda = b \tag{5}$$

---

*This paper will appear in Mathematical Programming and Its Applications, published by the National Institute of Higher Mathematics, Rome University.

for $\lambda = (\lambda_1, \ldots, \lambda_m)$, we obtain

$$\lambda \geq 0 \quad , \tag{6}$$

and thus the weighted sum of the x^i,

$$\hat{x}^o = \Sigma\, x^i \lambda_i \quad , \tag{7}$$

is starting feasible solution, i.e. satisfies (2).

One way to obtain the initial set of columns P_i for B is to solve the linear program

$$\begin{aligned} &\text{Min } cx \\ &\quad Ax = b \quad , \quad x \geq 0 \quad . \end{aligned} \tag{8}$$

The columns of the final basis of the linear program (or of any feasible basis generated by the simplex method) can be used to define the initial B where the i-th column in this case is some column j_i of A and thus x^i has all zero components except component j_i is unity. [Of course, if there exists no feasible solution to (8), then there would exist none for (2) and the procedure would terminate at this point.]

Algorithm

The iterative procedure on cycle t generates the linear program.

Find Min Z and $\lambda_j \geq 0$ such that

$$\begin{aligned} \gamma_1\lambda_1 + \gamma_2\lambda_2 + \cdots + \gamma_t\lambda_t &= Z \\ P_1\lambda_1 + P_2\lambda_2 + \cdots + P_t\lambda_t &= b \quad , \end{aligned} \tag{9}$$

where

$$P_j = Ax^j \text{ and } \gamma_j = F(x^j) \quad , \qquad \text{for some } x^j \geq 0 \quad . \tag{10}$$

Each cycle augments the linear program by one more column. On the initial cycle, we include, beside the columns corresponding to B, all the remaining columns of A corresponding to the linear terms of F(x), namely the p columns $A_{\cdot j}$ for $j > k$ not in B; for these p columns $P_i = A_{\cdot j}$, $\gamma_i = c_j$. Thus on the initial cycle there are $m + p$ columns. Instead of starting with $t = 0$, it simplifies subsequent notation if we formally call this initial cycle cycle $t = m + p$.

Let $\lambda_j = \lambda_j^t$ for $j = (1,\ldots,t)$ be the optimal basic feasible solution to the linear program (9). It is easy to see that the weighted sum

$$\hat{x}^t = \sum_{j=1}^{t} x^j \lambda_j^t \tag{11}$$

is a feasible solution of (2). Denote the indices of the columns forming the optimal feasible basis B^t of (9) by $(j_1, j_2, \ldots, j_m)$ which, of course, depend on t and let

$$\pi^t = (\gamma_{j_1}, \gamma_{j_2}, \ldots, \gamma_{j_m})[B^t]^{-1} \tag{12}$$

$$c^t = \bar{c} - \pi^t A \quad . \tag{13}$$

<u>Subproblem</u>

To find an improved solution, when such exists, we determine an $x = x^{t+1}$ which minimizes

$$\bar{F}(x) = F(x) - \pi^t Ax$$
$$= \bar{c}^t x + \sqrt{x_1^2 + x_2^2 + \cdots + x_k^2} \quad , \tag{14}$$

subject to $x_j \geq 0$ for $j = (1,\ldots,n)$ and the normalization condition

$$\sum_{j=1}^{k} x_j^2 = 1 \quad , \qquad (x_j \geq 0) \quad , \tag{15}$$

where

$$\bar{c}_j^t \geq 0 \qquad \text{for } j > k \tag{15a}$$

holds because of the inclusion of all columns $\begin{bmatrix} c_j \\ A_{\cdot j} \end{bmatrix}$ for all $j > k$ in the definition of (9). It follows by the duality theorem that an optimal solution to (9) will yield a π^t such that

$$c_j - \pi^t A_{\cdot j} = \bar{c}_j^t \geq 0 \qquad \text{for all } j > k \quad .$$

Assuming some $\bar{c}_j^t < 0$ for $j \leq k$, the minimum for the sub-problem (14), (15) is obtained by setting

$$x_j^{t+1} = 0 \qquad \text{for all } c_j^t \geq 0$$

and

$$x_j^{t+1} = -\bar{c}_j^t \Big/ \left[\sum_k (\bar{c}_k^t)^2\right]^{\frac{1}{2}} \qquad \text{for } \bar{c}_j^t < 0, \ \bar{c}_k^t < 0 \quad .$$

If all $\bar{c}_j^t \geq 0$ for all j, the algorithm is terminated (see below).

<u>Optimality Criterion</u>

The algorithm is terminated if $\bar{F}(x) = F(x) - \pi^t Ax \geq 0$ for all $x \geq 0$ [which, because $F(x)$ is homogeneous, is the same as for all x satisfying (15)]. The reason is that any feasible solution $\hat{x}$ would then satisfy $A\hat{x} = b$, $\hat{x} \geq 0$ so that $F(\hat{x}) \geq \pi^t b$. On the other hand, it is easy to show that the optimal basic feasible solution to (9)--namely $\hat{x}^t$ given by (11)--satisfies

$$\pi^t b = \Sigma\ F(x^i)\lambda_i^t \geq F(\hat{x}^t) \quad , \tag{16}$$

where the inequality follows from the convexity and homogeneity of F. We would then conclude that $F(\hat{x}) \geq F(\hat{x}^t)$ and thus $\hat{x}^t$ would be an optimal feasible solution to (1) and (2).

The new column for cycle t+1 is

$$\gamma_{t+1} = F(x^{t+1}),\ P_{t+1} = Ax^{t+1}$$

which "prices out" using the optimal basis of cycle t to be

$$\gamma_{t+1} - \pi^t P_{t+1} = F(x^{t+1}) - \pi^t A x^{t+1} = \overline{F}(x^{t+1}) \quad . \tag{17}$$

Hence no improvement to (9) can take place if $\overline{F}(x^{t+1}) \geq 0$; morever we have just shown that $\overline{F}(x^{t+1}) \geq 0$ also implies $\hat{x}^t$ optimal for (1) and (2). Accordingly, in this case we terminate. Otherwise we re-optimize (9) with t replaced by t+1.

Convergence

A formal proof of convergence to an optimal solution, which we now present, depends on $\{x|Ax = b,\ x \geq 0\}$ being bounded and assumes at least one basic solution being non-degenerate, i.e. $B^t\lambda = b$ solves with $\lambda > 0$ for some t. It also requires that all columns $\begin{pmatrix}\gamma_j \\ P_j\end{pmatrix}$ be kept as part of the linear program no matter how large t becomes. Convergence can be shown in the sense that $F(\hat{x}^t)$ converges to the finite minimum.

Note first that finite lower and upper bounds for π_i^t can be established for each i. For this purpose, it is convenient to let the non-degenerate basic solution be associated with the first m columns $B = [P_1, P_2, \ldots, P_m]$ and that (9) be multiplied by B^{-1} so that now the first m columns form an identity. The new right hand side b will now have all positive components (since the basic feasible solution associated with B is non-degenerate). Now by (16) $\pi^t b \geq F(\hat{x}^t) \geq \ell$ where ℓ is some finite lower bound for F(x). (Note that ℓ is finite because $\{x | Ax = b, x \geq 0\}$ is bounded.) Since π^t are optimal multipliers for (9), we have $\pi^t P_j \leq \gamma_j$ and hence $\pi_i^t \leq \gamma_i$ for $i = (1, \ldots, m)$. From this, it follows that

$$(\ell - \sum_{i \neq k} \gamma_i b_i)/b_k \leq \pi_k^t \leq \gamma_k \quad , \qquad b_i > 0,\ b_k > 0 \quad . \tag{18}$$

With π^t bounded, there exists a convergent subsequence S such that $\pi^t \to \pi^*$ for $t \in S$. On this subsequence S we can find a sub-subsequence S' for which $x^{t+1} \to x^* \neq 0$ because there is a point of condensation x^* on the hypersphere

$$\sum_{i=1}^{k} x_i^2 = 1,\ x_j = 0 \qquad \text{for } j > k \quad .$$

For $t \in S'$ we have

$$P^{t+1} \to P^* = Ax^* \quad ,$$

$$\gamma^{t+1} \to \gamma^* = F(x^*) \quad ,$$

and

$$\pi^t \to \pi^* \quad .$$

Let us choose $r \varepsilon S'$, $t \varepsilon S'$, $r < t$. Then

$$-\varepsilon^t = \gamma^{t+1} - \pi^t P^{t+1} < 0$$

and

$$\eta^r = \gamma^{r+1} - \pi^t P^{r+1} \geq 0 \qquad \text{for all } r < t$$

because π^t is the optimal vector of prices for the linear program for cycle t but not t+1. We now let $r \to \infty$, $t \to \infty$. Then

$$\eta^r + \varepsilon^t = (\gamma^{r+1} - \gamma^{t+1}) - \pi^t(P^{r+1} - P^{t+1}) \to 0 \quad ,$$

and thus

$$\eta^r \to 0, \ \varepsilon^t \to 0 \quad .$$

Let $\hat{x}$ be an optimal solution (1) and (2). Assuming there is at least one component $\hat{x}_j > 0$ for $j \leq k$, let $\beta = (\sum_{j=1}^{k} \hat{x}_j^{\,2})^{\frac{1}{2}}$ and $\overline{x} = \hat{x}/\beta$ be the "normalized" form of $\hat{x}$, then

$$\begin{aligned} -\varepsilon^t &= \gamma^{t+1} - \pi^t P^{t+1} \leq F(\overline{x}) - \pi^t A\overline{x} \\ -\varepsilon^t &\leq [F(\hat{x}) - \pi^t A\hat{x}]/\beta = [F(\hat{x}) - \pi^t b]/\beta \quad . \end{aligned} \tag{19}$$

If the assumption made above does not hold so that all $\hat{x}_j = 0$ for $j = (1,\ldots,k)$, then, because $\overline{c}_j^t \geq 0$ for $j > k$,

$$0 \leq F(\hat{x}) - \pi^t A\hat{x} = F(\hat{x}) - \pi^t b \quad . \tag{20}$$

It follows in either case that $F(\hat{x}) \geq \text{Lim}\ \pi^t b$ where the latter limit exists because $\pi^t b$ is monotonically decreasing and is bounded below by $\pi^{t+1} b \geq F(\hat{x}^{t+1}) \geq F(\hat{x})$ so that $F(\hat{x}) = \text{Lim}\ F(\hat{x}^{t+1})$.

General Comments

The procedure outlined can be expected to be efficient because it is the analogue of the simplex method. The method is similar to one proposed (without proof of convergence) for the chemical equilibrium problem and found on limited experimentation to have good convergence rates. Standard gradient procedures are not recommended because convergence would be too slow to be practical.

The method of proof presented here is along the lines first proposed by the author for convex functions which have a finite minimum for the subproblem without the normalization condition. By introducing minimization under the normalization condition (15), we have obtained a proof of convergence when $F(x)$ is convex and homogeneous of degree 1. For example, the proof given here is equally applicable to the chemical equilibrium problem. No bounds for $F(\hat{x})$, however, are given except by (19) and (16)

$$\pi^t b - \beta \varepsilon^t \leq F(\hat{x}) \leq \pi^t b \quad , \tag{21}$$

where unfortunately $\beta = \sqrt{\Sigma\ \hat{x}_j^2}$ for $j \leq k$ is not known.

An alternative method is to solve the problem by a parametric quadratic programming scheme where $cx = \theta$ is the unknown parameter and $\Sigma\ x_j^2$ for $j \leq k$ is minimized.

References

[1] Rosanov, Youri A. "A Few Methodological Remarks on Optimization of Random Cost Functions," IIASA Research Memorandum RM-73-7, December 1973.