数理解析研究所講究録
第 789 巻 1992 年 51-65

Application of the Alternating Direction Method of Multipliers to Separable Convex Programming Problems[1]

Masao FUKUSHIMA（福島　雅夫）

Department of Applied Mathematics and Physics
Faculty of Engineering, Kyoto University
Kyoto 606, Japan

Abstract: This paper presents a decomposition algorithm for solving convex programming problems with separable structure. The algorithm is obtained through application of the alternating direction method of multipliers to the dual of the convex programming problem to be solved. In particular, the algorithm reduces to the ordinary method of multipliers when the problem is regarded as nonseparable. Under the assumption that both primal and dual problems have at least one solution and the solution set of the primal problem is bounded, global convergence of the algorithm is established.

Keywords: Convex programming, separable problems, decomposition, alternating direction method of multipliers, parallel algorithm.

[1]This work was supported in part by the Scientific Research Grant-in-Aid from the Ministry of Education, Science and Culture, Japan (No. 02680025).

1. Introduction

Decomposition of large-scale problems is a classical topic of optimization [5], still attracting serious attention of many researchers. In particular, recent advances of parallel computers have demanded efficient algorithms that can take full advantage of a certain separable structure the problem to be solved may have. (See [2] and the references cited therein.)

The purpose of this paper is to present a new decomposition algorithm for solving the separable convex programming problem

$$\begin{aligned} \text{(P)} \quad \text{minimize} \quad & \sum_{j=1}^{n} f_j(x_j) \\ \text{subject to} \quad & \sum_{j=1}^{n} c_{ij}(x_j) \leq 0, \quad i = 1, ..., m, \\ & x_j \in X_j \subset R^{d_j}, \quad j = 1, ..., n, \end{aligned}$$

where $f_j : R^{d_j} \to R$ and $c_{ij} : R^{d_j} \to R$ are convex functions and X_j are nonempty closed convex subsets of R^{d_j} for all $i = 1, ..., m$ and $j = 1, ..., n$. In the following, all vectors are understood to be column vectors, but we shall often write as $x = (x_1, ..., x_n)$ instead of $x = (x_1, ..., x_n)^T$ or $x = (x_1^T, ..., x_n^T)^T$ in order to simplify the notation.

There are numbers of approaches to the solution of problem (P), but dual methods seem to be most popular among others. Let $y \in R^m$ be a vector of Lagrange multipliers and define the Lagrangian $L : R^d \times R^m \to R$ by

$$\begin{aligned} L(x, y) &= \sum_{j=1}^{n} f_j(x_j) + \langle y, \sum_{j=1}^{n} c_j(x_j) \rangle \\ &= \sum_{j=1}^{n} \{ f_j(x_j) + \langle y, c_j(x_j) \rangle \}, \end{aligned} \tag{1.1}$$

where $d = \sum_{j=1}^{n} d_j$, $x = (x_1, ..., x_n)$, $c_j(x_j) = (c_{1j}(x_j), ..., c_{mj}(x_j))$, $j = 1, ..., n$, and $\langle \cdot, \cdot \rangle$ denotes the inner product. Then the Lagrangian dual of (P) is the problem

$$\text{maximize} \quad g(y) \quad \text{subject to} \quad y \geq 0,$$

where the function $g : R^m \to [-\infty, +\infty)$ is defined by

$$g(y) = \inf_{x \in X} L(x, y) \tag{1.2}$$

with $X = X_1 \times \cdots \times X_n \subset R^d$. In the separable case, by (1.1) and (1.2), we may rewrite the dual problem as

$$\text{(D)} \quad \text{maximize} \quad \sum_{j=1}^{n} g_j(y) \quad \text{subject to} \quad y \geq 0,$$

where $g_j : R^m \to [-\infty, +\infty)$ are defined by

$$g_j(y) = \inf_{x_j \in X_j} \{f_j(x_j) + \langle y, c_j(x_j) \rangle\}, \quad j = 1, ..., n. \tag{1.3}$$

Thus the evaluation of the function g decomposes into evaluations of the n functions g_j, which can be done in parallel by solving n independent minimization problems involving each individual variable x_j only.

The dual problem (D) is a concave maximization problem. Moreover, if the infimum on the right-hand side of (1.3) is always attained uniquely for each j, which is particularly the case if the functions f_j are strictly convex and co-finite in the sense of [7, p. 116], then the dual functions g_j are not only finite-valued everywhere but also continuously differentiable, so that various descent methods can be applied to problem (D). Detailed discussions on dual descent methods for the case of linear constraints may be found in Tseng [12]. In the general case, however, the functions g_j are not necessarily differentiable, and further, it is quite likely that $g_j(y)$ may take the value $-\infty$ somewhere. A natural approach to such a problem would therefore be to use a carefully designed nondifferentiable optimization technique [6].

Another interesting way of dealing with problem (P) under the general setting is to modify the problem by adding a quadratic term to the objective function, thereby obtaining a strictly convex objective function. A typical example is the proximal point method [8, 9], of which each iteration consists of solving a subproblem of the form (P) with objective function replaced by $\sum_{j=1}^{n}\{f_j(x_j) + (r/2)||x_j - x_j^{(k)}||^2\}$, where $r > 0$ is a given constant and $x_j^{(k)}$ are components of the current iterate $x^{(k)}$. Since this problem has a strongly convex objective function, its dual is a differentiable concave maximization problem, to which various descent type algorithms can be applied.

A different but closely related approach is to utilize the method of multipliers [1]. Though straightforward application of the latter method to problem (P) generally loses the separable structure of the problem, careful reformulation of the problem may still lead to implementation that preserves the inherent separability for some special classes of problems. Specifically, Bertsekas and Tsitsiklis [2, pp. 249-251] consider the separable problem with linear constraints and show how the method of multipliers can be applied without destroying the separable structure of the given problem. Moreover, in [2, p. 254], it is shown that the same class of problems can also be dealt with effectively by the alternating direction method of multipliers [3, 4], which may be viewed as a variant of the method of multipliers. (See also [13, 14] for related methods.) Note that those methods do not require the strict convexity of the functions f_j, but assume that the coupling constraints are all linear.

In this paper, we consider applying the alternating direction method of multipliers to the dual problem (D), rather than the primal problem (P) as is done in [2]. The objective functions f_j are not assumed strictly convex and the constraint functions c_j are not assumed affine. Of course, none of the functions are supposed to be differentiable. Interestingly, the resulting algorithm resembles the method of Spingarn [11] that is derived from a variant of the proximal point method [10]. The difference between the Spingarn's algorithm and the present one might well be compared to that between the proximal method of multipliers [9] and the method of multipliers [1].

2. Preliminaries

In this section, we briefly review the alternating direction method of multipliers. For more detail, the reader may refer to [2, 3, 4].

The method is designed to solve a problem of the form

$$\begin{aligned} &\text{minimize} \quad G_1(y) + G_2(z) \\ &\text{subject to} \quad Ay - z = 0, \ \ y \in C_1, \ \ z \in C_2, \end{aligned} \tag{2.1}$$

where $G_1 : R^s \to (-\infty, +\infty]$ and $G_2 : R^t \to (-\infty, +\infty]$ are closed proper convex functions, A is a $t \times s$ matrix, and C_1 and C_2 are nonempty closed convex subsets of R^s and R^t, respectively. The iteration of the alternating direction method of multipliers may be written as

$$y^{(k+1)} := \arg\min_{y \in C_1} \{G_1(y) + \langle p^{(k)}, Ay \rangle + \frac{r}{2}||Ay - z^{(k)}||^2\}, \tag{2.2}$$

$$z^{(k+1)} := \arg\min_{z \in C_2} \{G_2(z) - \langle p^{(k)}, z \rangle + \frac{r}{2}||Ay^{(k+1)} - z||^2\}, \tag{2.3}$$

$$p^{(k+1)} := p^{(k)} + r(Ay^{(k+1)} - z^{(k+1)}), \tag{2.4}$$

where r is a positive constant and the initial vectors $p^{(0)}$ and $z^{(0)}$ may be chosen arbitrarily. Note that (2.2) and (2.3) correspond to a single cycle of the (block) Gauss-Seidel method to minimize the augmented Lagrangian

$$\Lambda_r(y, z, p^{(k)}) = G_1(y) + G_2(z) + \langle p^{(k)}, Ay - z \rangle + \frac{r}{2}||Ay - z||^2$$

for problem (2.1), while (2.4) is the ordinary multiplier update in the method of multipliers. The minimum on the right-hand side of (2.2) is uniquely attained whenever rank$(A) = s$, while the minimum on the right-hand side of (2.3) is always attained uniquely. Therefore the above method is well defined under the assumption rank$(A) = s$. Moreover, under the same assumption, it can be shown that the sequence $\{(y^{(k)}, z^{(k)}, p^{(k)})\}$ generated by (2.2)-(2.4) is bounded and every limit point of $\{(y^{(k)}, z^{(k)})\}$ is a solution of problem (2.1), whenever the solution set of the latter problem is nonempty. In addition, the sequence $\{(z^{(k)}, p^{(k)})\}$ has a unique limit point. (For a proof of these results, see Proposition 4.2 and its proof in [2, Chapter 3].)

3. Algorithm

In this section, we show how the alternating direction method of multipliers applied to the dual problem (D) yields a decomposition algorithm for solving problem (P). Throughout this section, we make the following assumption:

Assumption. Problems (P) and (D) have nonempty solution sets. Moreover, the solution set of (P) is bounded.

In applying the alternating direction method of multipliers to (D), we adopt a technique used in Bertsekas and Tsitsiklis [2, p. 246 and p. 256]. First we rewrite (D) in the following equivalent form:

$$(\hat{\mathrm{D}}) \quad \text{maximize} \quad \sum_{j=1}^{n} g_j(z_j)$$
$$\text{subject to} \quad y - z_j = 0, \quad j = 1, \dots, n,$$
$$z_j \geq 0, \quad j = 1, \dots, n,$$

where $z_j \in R^m$, $j = 1, \dots, n$, are artificial variables. We then apply the alternating direction method of multipliers (2.2)-(2.4) to problem $(\hat{\mathrm{D}})$ with the following identifications:

$$G_1(y) = 0, \qquad C_1 = R^m,$$
$$A = \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix} \in R^{nm \times m}, \qquad z = (z_1, z_2, \dots, z_n) \in R^{nm}, \tag{3.1}$$
$$G_2(z) = -\textstyle\sum_{j=1}^{n} g_j(z_j), \qquad C_2 = \{z \in R^{nm} \mid z_j \geq 0,\, j = 1, \dots, n\}.$$

Partitioning the multiplier vector $p \in R^{nm}$ as

$$p = (p_1, p_2, \dots, p_n),$$

where $p_j \in R^m$, $j = 1, \dots, n$, we may write the alternating direction method of multipliers for $(\hat{\mathrm{D}})$ as follows:

$$y^{(k+1)} := \arg\min_{y \in R^m} \{\langle \sum_{j=1}^{n} p_j^{(k)}, y\rangle + \frac{r}{2} \sum_{j=1}^{n} ||y - z_j^{(k)}||^2\}, \tag{3.2}$$
$$z_j^{(k+1)} := \arg\max_{z_j \geq 0} \{g_j(z_j) + \langle p_j^{(k)}, z_j\rangle - \frac{r}{2} ||y^{(k+1)} - z_j||^2\}, \quad j = 1, \dots, n, \tag{3.3}$$
$$p_j^{(k+1)} := p_j^{(k)} + r(y^{(k+1)} - z_j^{(k+1)}), \quad j = 1, \dots, n, \tag{3.4}$$

where r is a positive constant and the initial vectors $p_j^{(0)}$, $j = 1, \dots, n$, and $z_j^{(0)}$, $j = 1, \dots, n$, may be chosen arbitrarily. Note that, by the separability of $(\hat{\mathrm{D}})$, the updates

(3.3)-(3.4) of variables $z = (z_1, ..., z_n)$ and $p = (p_1, ..., p_n)$ can be performed in parallel for $j = 1, ..., n$.

Now let us go into details of the computation. First observe that, by (3.2), $y^{(k+1)}$ is explicitly written as

$$y^{(k+1)} = \frac{1}{n}\sum_{j=1}^{n} z_j^{(k)} - \frac{1}{nr}\sum_{j=1}^{n} p_j^{(k)}.$$

Next we consider (3.3). By the definition (1.3) of g_j, we have

$$\begin{aligned} &\max_{z_j \geq 0}\{g_j(z_j) + \langle p_j^{(k)}, z_j\rangle - \frac{r}{2}||y^{(k+1)} - z_j||^2\} \\ = \;&\max_{z_j \geq 0}\inf_{x_j \in X_j}\{f_j(x_j) + \langle z_j, p_j + c_j(x_j)\rangle - \frac{r}{2}||y^{(k+1)} - z_j||^2\}. \end{aligned} \tag{3.5}$$

Under the standing assumption that (P) has a nonempty bounded solution set, we can show that $z_j^{(k+1)} = (z_{1j}^{(k+1)}, ..., z_{mj}^{(k+1)})$ is given by

$$z_{ij}^{(k+1)} = \max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j^{(k+1)}))\}, \quad i = 1, ..., m, \tag{3.6}$$

where $x_j^{(k+1)}$ is a solution of the minimization problem

$$\begin{aligned} &\text{minimize} \quad f_j(x_j) + \frac{r}{2}\sum_{i=1}^{m}\left[\max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j))\}\right]^2 \\ &\text{subject to} \quad x_j \in X_j. \end{aligned} \tag{3.7}$$

To see this, first notice that if the solution set of problem (P) is nonempty and bounded, then the functions $\sum_{j=1}^{n} f_j$, $\sum_{j=1}^{n} c_{ij}$, $i = 1, ..., m$, and the set $X = X_1 \times \cdots \times X_n$ have no direction of recession in common in the sense of [7, p. 61 and p. 69]. By the separability of the problem, this implies that the same is true for the functions f_j, $c_{ij}, i = 1, ..., m$, and the set X_j for each j.

Let us consider the saddle function $K : R^{d_j} \times R^m \to [-\infty, +\infty]$ defined by

$$K(x_j, z_j) = \begin{cases} f_j(x_j) + \langle z_j, p_j^{(k)} + c_j(x_j)\rangle - \frac{r}{2}||y^{(k+1)} - z_j||^2, & \text{if } x_j \in X_j, z_j \geq 0 \\ +\infty, & \text{if } x_j \notin X_j, z_j \geq 0 \\ -\infty, & \text{if } z_j \ngeq 0. \end{cases}$$

Then, by the preceding arguments, it is seen that the convex function $K(\cdot, z_j)$ has no direction of recession for any $z_j > 0$, while the convex function $-K(x_j, \cdot)$ trivially has no direction of recession for any $x_j \in X_j$. Therefore it follows from [7, Theorems 37.3 and 37.6] that

$$\sup_{z_j \geq 0} \inf_{x_j \in X_j} K(x_j, z_j) = \inf_{x_j \in X_j} \sup_{z_j \geq 0} K(x_j, z_j) < \infty$$

and that the function K actually has a saddle point $(\bar{x}_j, \bar{z}_j) \in X_j \times \{z_j \mid z_j \geq 0\}$ such that

$$K(\bar{x}_j, \bar{z}_j) = \max_{z_j \geq 0} \min_{x_j \in X_j} K(x_j, z_j) = \min_{x_j \in X_j} \max_{z_j \geq 0} K(x_j, z_j). \tag{3.8}$$

Consequently, it follows from (3.5) that

$$\begin{aligned} &\max_{z_j \geq 0}\{g_j(z_j) + \langle p_j^{(k)}, z_j \rangle - \frac{r}{2}||y^{(k+1)} - z_j||^2\} \\ = \;&\min_{x_j \in X_j} \max_{z_j \geq 0}\{f_j(x_j) + \langle z_j, p_j + c_j(x_j) \rangle - \frac{r}{2}||y^{(k+1)} - z_j||^2\}. \end{aligned} \tag{3.9}$$

But, for any fixed x_j, the maximum on the right-hand side of (3.9) is uniquely attained by

$$z_j = \left[y^{(k+1)} + \frac{1}{r}(p_j^{(k)} + c_j(x_j))\right]_+,$$

where $[\cdot]_+$ denotes the orthogonal projection of a vector onto the nonnegative orthant, i.e.,

$$z_{ij} = \max\,\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j))\}, \quad i = 1, ..., m. \tag{3.10}$$

We may thus substitute (3.10) into the function on the right-hand side of (3.9) to eliminate the variables z_j. As a result, we obtain the objective function of problem (3.7). Therefore, if a solution $x_j^{(k+1)}$ of problem (3.7) is found, we can determine $z_j^{(k+1)}$ by (3.6). Clearly such $(x_j^{(k+1)}, z_j^{(k+1)})$ is a saddle point $(\bar{x}_j, \bar{z}_j)$ of K satisfying (3.8). (Note that the previous arguments guarantee the existence of a solution of problem (3.7).)

To summarize, we state the algorithm as follows:

Algorithm 1.

Step 1: Choose a constant $r > 0$ and initial vectors $(p_j^{(0)}, z_j^{(0)})$, $j = 1, ..., n$, arbitrarily. Set $k := 0$.

Step 2: Compute

$$y^{(k+1)} = \frac{1}{n}\sum_{j=1}^{n} z_j^{(k)} - \frac{1}{nr}\sum_{j=1}^{n} p_j^{(k)}. \tag{3.11}$$

Step 3: For each $j = 1, ..., n$, find a solution $x_j^{(k+1)}$ of the minimization problem

$$\begin{array}{ll} \text{minimize} & f_j(x_j) + \dfrac{r}{2}\displaystyle\sum_{i=1}^{m}\left[\max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j))\}\right]^2 \\ \text{subject to} & x_j \in X_j, \end{array}$$

and determine $z_j^{(k+1)} = (z_{1j}^{(k+1)}, ..., z_{mj}^{(k+1)})$ by

$$z_{ij}^{(k+1)} = \max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j^{(k+1)}))\}, \quad i = 1, ..., m. \tag{3.12}$$

Step 4: For each $j = 1, ..., n$, compute

$$p_j^{(k+1)} = p_j^{(k)} + r(y^{(k+1)} - z_j^{(k+1)}).$$

Set $k := k + 1$ and go to Step 2. □

Since the matrix A defined by (3.1) has full column rank, it follows from the fact mentioned in the previous section that the sequence $\{(y^{(k)}, z^{(k)}, p^{(k)})\}$ generated by Algorithm 1 is bounded. Moreover, every limit point of $\{(y^{(k)}, z^{(k)})\}$ is a solution of problem ($\hat{\mathrm{D}}$), and the sequence $\{(z^{(k)}, p^{(k)})\}$ has a unique limit point. It now remains to establish convergence of the sequence $\{x^{(k)}\}$.

Theorem 3.1 *Suppose that problems* (P) *and* (D) *have nonempty solution sets, and that the solution set of* (P) *is bounded. Then the sequence* $\{x^{(k)}\}$ *generated by Algorithm 1 is bounded and every limit point of* $\{x^{(k)}\}$ *is a solution of* (P).

Proof. Since $\{(z^{(k)}, p^{(k)})\}$ has a unique limit point, and since any limit point of $\{(y^{(k)}, z^{(k)})\}$ is a solution of problem $(\hat{\mathrm{D}})$, the sequence $\{y^{(k)}\}$ also has a unique limit point, which is equal to that of $\{z_j^{(k)}\}$ for any j. That is, we have

$$y^{(k)} \to y^*, \tag{3.13}$$

$$z_j^{(k)} \to y^*, \qquad j = 1, ..., n, \tag{3.14}$$

$$p_j^{(k)} \to p_j^*, \qquad j = 1, ..., n, \tag{3.15}$$

for some $y^* \in R^m$ and $p_j^* \in R^m$, $j = 1, ..., n$, where in particular y^* is a solution of problem (D).

For each j, let us define the functions $F_j^{(k)} : R^{d_j} \to (-\infty, +\infty]$, $k = 1, 2, \dots$, and $F_j^* : R^{d_j} \to (-\infty, +\infty]$ by

$$F_j^{(k)}(x_j) = f_j(x_j) + \frac{r}{2} \sum_{i=1}^{m} \left[\max\{0, y_i^{(k)} + \frac{1}{r}(p_{ij}^{(k-1)} + c_{ij}(x_j))\} \right]^2 + \delta(x_j | X_j)$$

and

$$F_j^*(x_j) = f_j(x_j) + \frac{r}{2} \sum_{i=1}^{m} \left[\max\{0, y_i^* + \frac{1}{r}(p_{ij}^* + c_{ij}(x_j))\} \right]^2 + \delta(x_j | X_j),$$

respectively, where $\delta(\cdot | X_j)$ is the indicator function of the set X_j. Note that the sequence $\{F_j^{(k)}\}$ of closed convex functions e-converges (epi-converges) to the closed convex function F_j^* in the sense of [15]. Moreover, since (P) has a nonempty bounded solution set, the functions f_j, c_{ij}, $i = 1, ..., m$, and the set X_j have no direction of recession in common, which in turn implies that the function F_j^* has no direction of recession and hence has a compact solution set. Since $x_j^{(k)}$ is a minimum of the function $F_j^{(k)}$ for each k, it follows from [15, Theorem 9] that the sequence $\{x_j^{(k)}\}$ is bounded and every limit point belongs to the set of minima of F_j^*.

Now let x_j^* denote an arbitrary limit point of $\{x_j^{(k)}\}$ for each j. Since x_j^* minimizes the function F_j^*, we have

$$\begin{aligned} 0 &\in \partial F_j^*(x_j^*) \\ &= \partial f_j(x_j^*) + \sum_{i=1}^{m} \max\{0, y_i^* + \frac{1}{r}(p_{ij}^* + c_{ij}(x_j^*))\} \partial c_{ij}(x_j^*) + \partial \delta(x_j^* | X_j), \end{aligned} \tag{3.16}$$

where ∂ denotes the subdifferential operator.

Incidentally it follows from (3.11) and (3.13)-(3.15) that

$$\sum_{j=1}^{n} p_j^* = 0. \tag{3.17}$$

Moreover, (3.12) together with (3.13)-(3.15) implies

$$y_i^* = \max\{0, y_i^* + \frac{1}{r}(p_{ij}^* + c_{ij}(x_j^*)\}, \quad i = 1, ..., m, \tag{3.18}$$

which in turn implies that

$$\begin{aligned} y_i^* = 0 \quad &\Rightarrow \quad p_{ij}^* + c_{ij}(x_j^*) \le 0, \\ y_i^* > 0 \quad &\Rightarrow \quad p_{ij}^* + c_{ij}(x_j^*) = 0. \end{aligned} \tag{3.19}$$

Then it follows from (3.16) and (3.18) that

$$0 \in \partial f_j(x_j^*) + \sum_{i=1}^{m} y_i^* \partial c_{ij}(x_j^*) + \partial \delta(x_j^* | X_j). \tag{3.20}$$

Since (3.20) holds for each j, we have

$$\begin{aligned} 0 \quad &\in \quad \sum_{j=1}^{n} \partial f_j(x_j^*) + \sum_{i=1}^{m} y_i^* \sum_{j=1}^{n} \partial c_{ij}(x_j^*) + \sum_{j=1}^{n} \partial \delta(x_j^* | X_j) \\ &= \quad \partial \left(\sum_{j=1}^{n} f_j(x_j^*) \right) + \sum_{i=1}^{m} y_i^* \partial \left(\sum_{j=1}^{n} c_{ij}(x_j^*) \right) + \partial \left(\sum_{j=1}^{n} \delta(x_j^* | X_j) \right), \end{aligned} \tag{3.21}$$

where the last equality follows from [7, Theorem 23.8]. On the other hand, the relation (3.19) implies that the inequalities

$$p_{ij}^* + c_{ij}(x_j^*) \le 0$$

hold for all i and j, so that

$$\sum_{j=1}^{n} p_{ij}^* + \sum_{j=1}^{n} c_{ij}(x_j^*) \le 0, \quad i = 1, ..., m.$$

Therefore, by (3.17), we have

$$\sum_{j=1}^{n} c_{ij}(x_j^*) \le 0, \quad i = 1, ..., m. \tag{3.22}$$

Moreover, by (3.19) and (3.17), we have

$$y_i^* > 0 \quad \Rightarrow \quad \sum_{j=1}^{n} c_{ij}(x_j^*) = 0. \tag{3.23}$$

Since (3.21)-(3.23) imply that $x^* = (x_1^*, ..., x_n^*)$ together with the Lagrange multiplier vector $p^* = (p_1^*, ..., p_m^*)$ satisfies the Karush-Kuhn-Tucker optimality conditions for problem (P), the proof is complete. □

4. Discussion

The algorithm presented in the previous section solves at each iteration the following n independent subproblems:

$$\begin{array}{ll} \text{minimize} & f_j(x_j) + \frac{r}{2}\sum_{i=1}^{m}\left[\max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j))\}\right]^2 \\ \text{subject to} & x_j \in X_j. \end{array} \tag{4.1}$$

We remark that the objective function of problem (4.1) looks like an augmented Lagrangian for the problem

$$\begin{array}{ll} \text{minimize} & f_j(x_j) \\ \text{subject to} & c_{ij}(x_j) \le -p_{ij}^{(k)}, \quad i = 1, ..., m, \\ & x_j \in X_j. \end{array}$$

with the Lagrange multiplier vector $y^{(k+1)} \in R^m$ that is common to all $j = 1, ..., n$. Therefore, the vector $p_j^{(k)} = (p_{1j}^{(k)}, ..., p_{mj}^{(k)})$ may be thought of as the (negative) amount of the resources assigned to the jth subsystem. After solving subproblems (4.1), the algorithm updates the Lagrange multiplier vector separately for each j based on the respective solutions of (4.1). At this stage, there are n different estimates $z_j^{(k+1)}$ of Lagrange multipliers for $j = 1, ..., n$. Using this information, the algorithm then reallocates the resources by updating $p_j^{(k+1)}$, and proceeds to the next iteration. At the beginning of the new iteration, the different values $z_j^{(k+1)}$ of Lagrange multiplier estimates are integrated to the single Lagrange multiplier vector $y^{(k+1)}$, which is again common to all subsystems. In this manner, the algorithm successively updates not only Lagrange multipliers (prices) for the coupling constraints but also the amount of resources to be assigned to each subsystem.

A decomposition method with similar nature has been proposed by Spingarn [11] for the same class of separable problems (P). Though the original notation and formulation of [11] are somewhat different from ours, suitable transformation of variables and rearrangement reveal that Spingarn's algorithm, which is called Algorithm 4 in [11], may be restated as follows:

Step 1: Choose a constant $r > 0$ and initial vectors $(z_j^{(0)}, p_j^{(0)})$, $j = 1, ..., n$, such that $\sum_{j=1}^n p_j^{(0)} = 0$. Set $k := 0$.

Step 2: Compute

$$y^{(k+1)} = \frac{1}{n} \sum_{j=1}^{n} z_j^{(k)}.$$

Step 3: For each $j = 1, ..., n$, find the unique solution $x_j^{(k+1)}$ of the minimization problem

$$\begin{array}{ll} \text{minimize} & f_j(x_j) + \frac{1}{2n^2 r} \|x_j - x_j^{(k)}\|^2 + \frac{r}{2} \sum_{i=1}^{m} \left[\max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j))\} \right]^2 \\ \text{subject to} & x_j \in X_j, \end{array}$$

and determine $z_j^{(k+1)} = (z_{1j}^{(k+1)}, ..., z_{mj}^{(k+1)})$ by

$$z_{ij}^{(k+1)} = \max\{0, y_i^{(k+1)} + \frac{1}{r}(p_{ij}^{(k)} + c_{ij}(x_j^{(k+1)}))\}, \quad i = 1, ..., m.$$

Step 4: For each $j = 1, ..., n$, compute

$$\hat{p}_j^{(k+1)} = p_j^{(k)} + r(y^{(k+1)} - z_j^{(k+1)})$$

and

$$p_j^{(k+1)} = \hat{p}_j^{(k+1)} - \frac{1}{n} \sum_{\ell=1}^{n} \hat{p}_\ell^{(k+1)}.$$

Set $k := k + 1$ and go to Step 2. □

This algorithm differs from Algorithm 1 in two respects. First, each subproblem solved at Step 3 contains the extra quadratic term $\frac{1}{2n^2r}||x_j - x_j^{(k)}||^2$, which is peculiar to methods of the proximal point type. Second, Step 4 contains an additional update of the multiplier vectors $p_j^{(k)}$ in order to maintain the condition $\sum_{j=1}^n p_j^{(k)} = 0$ throughout the iterations. In the special case where $n = 1$, Spingarn's algorithm reduces to the proximal method of multipliers [9], while Algorithm 1 is nothing but the ordinary method of multipliers [1].

References

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

[2] D.P. Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[3] M. Fortin and R. Glowinski, On decomposition-coordination methods using an augmented Lagrangian, in M. Fortin and R. Glowinski (eds.), *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, North-Holland, Amsterdam, 1983, pp. 97-146.

[4] D. Gabay and B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, *Computers and Mathematics with Applications* vol.**2**, pp. 17-40, 1976.

[5] L.S. Lasdon, *Optimization Theory for Large Systems*, Macmillan, New York, 1970.

[6] C. Lemaréchal, Nondifferentiable optimization, in G.L. Nemhauser, A.H.G. Rinnooy Kan and M.J. Todd (eds.), *Handbooks in Operations Research and Management Science, Vol. 1, Optimization*, North-Holland, Amsterdam, 1989, pp. 529-572.

[7] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[8] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization* vol.**14**, pp. 877-898, 1976.

[9] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Mathematics of Operations Research* vol.**1**, pp. 97-116, 1976.

[10] J.E. Spingarn, Partial inverse of a monotone operator, *Applied Mathematics and Optimization* vol.**10**, pp.247-265, 1983.

[11] J.E. Spingarn, Applications of the method of partial inverses to convex programming: Decomposition, *Mathematical Programming* vol.**32**, pp.199-223, 1985.

[12] P. Tseng, Dual ascent methods for problems with strictly convex costs and linear constraints: A unified approach, *SIAM Journal on Control and Optimization* vol.**28**, pp. 214-242, 1990.

[13] P. Tseng, Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming, *Mathematical Programming* vol.**48**, pp. 249-263, 1990.

[14] P. Tseng, Applications of a splitting algorithm to decomposition in convex programming and variational inequalities, *SIAM Journal on Control and Optimization* vol.**29**, pp. 119-138, 1991.

[15] R.J.B. Wets, Convergence of convex functions, variational inequalities and convex optimization problems, in R.W. Cottle, F. Giannessi and J.-L. Lions (eds.), *Variational Inequalities and Complementarity Problems*, John Wiley, Chichester, 1980, pp. 375-403.