The Implementation of the Primal-Dual Interior-Point Method for the Semidefinite Programs and its Engineering Applications

Katsuki Fujisawa

Department of Architecture and Architectural Systems, Kyoto University.

In recent years, semidefinite program (SDP) has been intensively studied both in theoretical and practical aspects of various fields including interior-point methods, combinatorial optimization and the control and systems theory. The SDPA (SemiDefinite Programming Algorithm) [4] is a C++ implementation of a Mehrotra-type primal-dual predictor-corrector interior-point method for solving the standard form semidefinite program. The SDPA incorporates data structures for handling sparse matrices and an efficient method proposed by Fujisawa et al. [5] for computing search directions when problems to be solved are large scale and sparse. Finally, we report numerical experiments of the SDP for the structural optimization under multiple eigenvalue constraints.

1 Introduction.

The main purpose of this paper is to explain the implementation of SDPA (SemiDefinite Programming Algorithm) [4] for semidefinite programs and report some numerical experiments of SDPA. Besides SDPA, there are some computer programs SDPpack [2], SDPSOL [19], CSDP [3], SDPHA [15] and SDPT3 [17] for semidefinite programs which are available through the Internet. The SDPA is a C++ implementation of a Mehrotra-type [10] primal-dual predictor-corrector interior-point method. Although three types of search directions, the HRVW/KSH/M direction [9], the NT direction [13] and the AHO direction [1] are available in SDPA, we employed the HRVW/KSH/M direction in our numerical experiments because its computation is the cheapest among the three directions (particularly, for sparse data matrices) when we employ the method proposed by Fujisawa et al. [5]. Monteiro et al. [12] recently showed that in theory, the NT direction requires less computation for dense matrices. However, their method needs large amount of computational memory and does not efficiently exploit the sparse data structures. Actually, according to their numerical results, the computation of the HRVW/KSH/M direction is favorable compared to the computation of the NT and AHO directions.

Let $\mathbb{R}^{n \times n}$ and $\mathcal{S}_n \subset \mathbb{R}^{n \times n}$ denote the set of all $n \times n$ real matrices and the set of all $n \times n$ real symmetric matrices, respectively. We use the notation $U \cdot V$ for the inner product of $U$, $V \in \mathbb{R}^{n \times n}$, i.e., $U \cdot V = \sum_{i=1}^{n} \sum_{j=1}^{n} U_{ij} V_{ij}$, where $U_{ij}$ and $V_{ij}$ denote the ($i$, $j$)th element of $U$ and $V$, respectively. We write $X \succeq O$ and $X \succ O$ when $X \in \mathcal{S}_n$ is positive semidefinite and positive definite, respectively.

Let $A_i \in \mathcal{S}_n$ ($0 \leq i \leq m$) and $b_i \in \mathbb{R}$ ($1 \leq i \leq m$). SDPA solves the standard form semidefinite program and its dual:

\[
\begin{align*}
\mathcal{P}: \text{minimize} & \quad A_0 \cdot X \\
\text{subject to} & \quad A_i \cdot X = b_i \ (1 \leq i \leq m), \quad X \succeq O.
\end{align*}
\]

\[
\begin{align*}
\mathcal{D}: \text{maximize} & \quad \sum_{i=1}^{m} b_i y_i \\
\text{subject to} & \quad \sum_{i=1}^{m} A_i y_i + Z = A_0, \quad Z \succeq O.
\end{align*}
\]

For simplicity, we say that $(X, y, Z)$ is a feasible solution (an interior-feasible solution, or an optimal solution, respectively) of the SDP (1) if $X$ is a feasible solution (an interior-feasible solution, i.e., a feasible solution satisfying $X \succ O$ or a minimizing solution, respectively) of $\mathcal{P}$ and $(y, Z)$ is a feasible solution (an interior-feasible solution, i.e., a feasible solution satisfying $Z \succ O$ or a maximizing solution, respectively) of $\mathcal{D}$.

In Section 2, we present some issues on the implementation of SDPA which are relevant for our numerical experiments. Section 3 is devoted to SDPs for structural optimization under multiple
eigenvalue constraints. In Section 4, we present numerical results of the SDPA. Section 5 gives conclusions.

2 Some Issues of the Implementation of SDPA.

In this section, we describe the data structures for handling block diagonal matrices and the HRVW/KSH/M search direction both used in SDPA. Finally, we explain the algorithmic framework of SDPA.

2.1 Data Structures

The SDPA can handle block diagonal matrices. In terms of the number of blocks, denoted by nBLOCK, and the block structure vector, denoted by bLOCKsTRUCT, we express a common matrix data structure for the constraint matrices $A_0, A_1, \ldots, A_m$. If we deal with a block diagonal matrix $A$ of the form

$$A = \begin{pmatrix} B_1 & O & O & \cdots & O \\ O & B_2 & O & \cdots & O \\ . & . & . & \cdots & . \\ O & O & O & \cdots & B_\ell \end{pmatrix},$$

\[B_i: \text{a } p_i \times p_i \text{ symmetric matrix } (i = 1, 2, \ldots, \ell),\]

we define the number nBLOCK of blocks and the block structure vector bLOCKsTRUCT as follows:

nBLOCK = $\ell$,

bLOCKsTRUCT = ($\beta_1$, $\beta_2$, $\ldots$, $\beta_\ell$),

$\beta_i = \begin{cases} p_i & \text{if } B_i \text{ is a symmetric matrix}, \\ -p_i & \text{if } B_i \text{ is a diagonal matrix}. \end{cases}$

For example, if $A$ is of the form

$$\begin{pmatrix} 1 & 2 & 3 & 0 & 0 & 0 \\ 2 & 4 & 5 & 0 & 0 & 0 \\ 3 & 5 & 6 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix},$$

we have

nBLOCK = 3 and bLOCKsTRUCT = (3, 2, -2).

2.2 Search Direction.

The HRVW/KSH/M direction is the solution $(dX, dy, dZ)$ of the system of equations

$$A_i \cdot dX = p_i \ (1 \leq i \leq m), \ dX \in S^n, \quad (4)$$

$$\sum_{i=1}^{m} A_i dy_i + dZ = D, \ dZ \in S^n, \quad (5)$$

$$\hat{dX} Z + X dZ = K, \ \hat{dX} \in R^{n \times n}, \ dX = (\hat{dX} + \hat{dX}^T) / 2. \quad (6)$$
Here \( p_i \in R, D \in S^n \) and \( K \in R^{n \times n} \) denote a scalar constant, an \( n \times n \) constant symmetric matrix, and an \( n \times n \) constant matrix, respectively, which are determined by the current point \((X, y, Z)\) and some other factors. Note that \( \hat{X} \in R^{n \times n} \) serves as an auxiliary variable matrix. Under the linear independence assumption on the set \( \{A_i : 1 \leq i \leq m\} \) of constraint matrices, we know [9] that for any \( X \succ O, Z \succ O, p_i \in R \) \( (1 \leq i \leq m) \), \( D \in S^n \), and \( K \in R^{n \times n} \), the system of equations (4), (5) and (6) has a unique solution \((dX, dy, dZ)\).

We can reduce the system of equations (4), (5) and (6) to

\[
Bdy = b, \\
dZ = D - \sum_{i=1}^{m} A_idy_i, \\
\hat{X} = (K - XDZ)^{-1}, \quad \hat{X} = (\hat{X} + \hat{X}^T)/2, \\
\]

where

\[
B_{ij} = XA_iZ^{-1} \cdot A_j \quad (1 \leq i \leq m, \ 1 \leq j \leq m) \\
b_i = p_i - (K - XDZ)^{-1} \cdot A_i \quad (1 \leq i \leq m).
\]

The matrices \( X, Z^{-1} \) and \( B \) are symmetric and dense in general even when all \( A_i \) \( (1 \leq i \leq m) \) are sparse. Hence solving the system of equations (7) in \( dy \) by using a direct method such as the Cholesky factorization and the LDL\(^T\) factorization of \( B \) requires \( O(m^3) \) arithmetic operations. On the other hand, if we treat all \( A_i \) \( (1 \leq i \leq m) \) as dense matrices and if we use the above formulae (9) for the coefficient matrix \( B \) in a straightforward way, the computation of \( B \) requires \( O(mn^3 + mn^2) \) arithmetic operations. Therefore computing the coefficient matrix \( B \) is more crucial than solving \( Bdy = b \) in the entire computation of the HRVW/KSH/M direction.

In their paper [5], Fujisawa, Kojima and Nakata proposed three distinct formulae \( F_1, F_2 \) and \( F_3 \) for computing \( B \), and their efficient combination \( F_* \). They demonstrated through numerical experiments that the combined formula \( F_* \) worked very efficiently when some of \( A_i \) \( (1 \leq i \leq m) \) are sparse. We incorporated their formula \( F_* \) into SDPA. See the paper [5] for more details.

### 2.3 The Algorithmic Framework of SDPA.

**Step 0:** Set an initial point \((X^0, y^0, Z^0)\) with \( X^0 \succ O, Z^0 \succ O \). Decide on the search direction to use. Set the parameters: \( 0.0 < \epsilon^*, 1 < \omega^*, 0.01 \leq \beta^* \leq 0.1 \) and \( \beta^* \leq \bar{\beta} \leq 0.2 \). (The default values of these parameters are: \( \epsilon^* = 1.0E-5, \omega^* = 2.0, \beta^* = 0.05 \) and \( \bar{\beta} = 0.1 \).) Let \( k = 0 \).

**Step 1:** If the current iterate \((X^k, y^k, Z^k)\) is feasible and the relative gap

\[
\frac{|P - D|}{\max \{1.0, (|P| + |D|)/2\}}
\]

gets smaller than \( \epsilon^* \), then stop the iteration. Here \( P \) and \( D \) denote the primal and the dual objective values, respectively. If we detect that there is no feasible solution \((X, y, Z)\) such that \( \omega^*X^0 \succeq X \succeq O \) and \( \omega^*Z^0 \succeq Z \succeq O \), then stop the iteration. See [9] for details on how to get such information on infeasibility.

**Step 2:** (Predictor Step) Let

\[
\beta_p = \begin{cases}
0 & \text{if the current iterate is feasible}, \\
\bar{\beta} & \text{otherwise}.
\end{cases}
\]

Solve the system of equations (4), (5) and (6) with \( K = \beta_p(X \bullet Z/n)I - XZ \) to compute the predictor direction \((dX_p, dy_p, dZ_p)\).
Step 3: (Corrector Step) Let
\[ \beta = \frac{(X + \bar{\alpha}_p dX_p) \bullet (Z + \bar{\alpha}_d dZ_p)}{(X \bullet Z)}, \]
where \(\bar{\alpha}_p\) and \(\bar{\alpha}_d\) are computed as in [6]. Choose the parameter \(\beta_c\) as follows:
\[ \beta_c = \begin{cases} 
\max\{\beta^*, \beta^2\} & \text{if the current iterate is feasible and } \beta \leq 1.0, \\
\max\{\beta, \beta^2\} & \text{if the current iterate is infeasible and } \beta \leq 1.0, \\
1.0 & \text{otherwise.} 
\end{cases} \]
Compute the corrector direction \((dX_c, dy_c, dZ_c)\) by solving the system of equations (4), (5) and (6) with \(K = \beta_c(X \bullet Z/n)I - XZ - dX_p dZ_p\).

Step 4: Set the next iterate \((X^{k+1}, y^{k+1}, Z^{k+1})\) such that
\[ X^{k+1} = X^k + \alpha_p dX_c \quad \text{and} \quad (y^{k+1}, Z^{k+1}) = (y^k, Z^k) + \alpha_d (dy_c, dZ_c), \]
where \(\alpha_p\) and \(\alpha_d\) are computed as in [6].

Step 5: \(k \leftarrow k + 1\) and go to Step 1.

3 Semidefinite Programming for Structural Optimization under Multiple Eigenvalue Constraints.

The eigenvalues of free vibration as well as the linear buckling load factor, which are calculated by solving linear eigenvalue problems, are important performance measures of the structures. Therefore, there have been many researches for optimization of structures under eigenvalue constraints. Consider a truss with fixed locations of nodes and members which can exist. The vector of member cross-sectional areas is denoted by \(A = \{A_i\}\). Let \(K\) and \(M_s\) denote the stiffness matrix and the mass matrix due to the structural mass which are functions of \(A\). The mass matrix for nonstructural mass is denoted by \(M_0\). The eigenvalue problem of vibration is formulated as
\[ K\Phi_r = \Omega_r(M_s + M_0)\Phi_r, \quad (r = 1, 2, \ldots, N^d) \] (10)
where \(\Omega_r\) and \(\Phi_r\) are the \(r\)th eigenvalue and eigenvector, and \(N^d\) is the number of freedom of displacements. The eigenvector \(\Phi_r\) is normalized by
\[ \Phi_r^T M \Phi_r = 1. \quad (r = 1, 2, \ldots, N^d) \] (11)

Let \(\bar{\Omega}\) denote the specified lower bound of the eigenvalues. The topology optimization problem for specified fundamental eigenvalue is formulated as
\[ \text{OPE: Minimize} \sum_{i=1}^{N^m} L_i A_i \] (12)
subject to \(\Omega_r \geq \bar{\Omega}, \quad (r = 1, 2, \ldots, N^d) \) (13)
\[ A_i \geq 0, \quad (i = 1, 2, \ldots, N^m) \] (14)
where \(L_i\) is the length of the \(i\)th member, and \(N^m\) is the number of members. The optimal topology is found by removing the members with \(A_i \leq \epsilon\), where \(\epsilon\) is a small positive lower bound. A small positive lower bound is usually given for \(A_i\) to prevent instability of the structure.

It is well known that optimum designs for specified fundamental eigenvalue often have multiple or repeated eigenvalues. If the fundamental eigenvalue of the optimum design is simple, OPE
may easily be solved by using a nonlinear programming or an optimality criteria approach [18, 16], because there is no difficulty in calculating the sensitivity coefficients of Ω₁ with respect to Aᵢ. In the case of multiple eigenvalue, only directional sensitivity coefficients can be calculated [7]. Although some formulations of sensitivity analysis of repeated eigenvalue has been presented [8, 11], it is not clear if those formulations can be used for optimizing large structures. In this paper, the eigenvalue problem (10) is converted into a standard form, and solved by using semi-definite programming.

Consider a structure where Ω₁ ≥ ¯Ω is satisfied. In this case the Rayleigh’s principle leads to the following inequality for any admissible mode ψ:

$$\psi^T(K - \bar{\Omega}M)\psi \geq 0$$  \hfill (15)

where (11) has been used. This inequality implies that the matrix K − ΩM is positive semi-definite, and formulations of SDP may be possible. The matrices Kᵢ and Mᵢ are defined as

$$K_i = \frac{\partial K}{\partial A_i}, \quad M_i = \frac{\partial M_s}{\partial A_i}.$$  \hfill (16)

Since K and Mₛ are linear functions of Aᵢ for trusses, those are written as

$$K = \sum_{i=1}^{N_m} K_i A_i, \quad M_s = \sum_{i=1}^{N_m} M_i A_i$$  \hfill (17)

The problems P and D for this case are formulated as

$$P':\ \text{Minimize} \sum_{i=1}^{N_m} L_i y_i$$ \hfill (18)

subject to \ $X = \sum_{i=1}^{N_m} (K_i - \bar{\Omega}M_i) y_i - \bar{\Omega}M_0$ \hfill (19)

\ $X \in S^n, \ X \succeq O.$ \hfill (20)

$$D':\ \text{Maximize} \ \bar{\Omega}M_0 \bullet Y$$ \hfill (21)

subject to \ $(K_i - \bar{\Omega}M_i) \bullet Y = L_i$ \hfill (22)

\ $Y \in S^n, \ Y \succeq O.$ \hfill (23)

Problems P' and D' are solved successively to find optimal solutions. It is important to note here that the design sensitivity coefficients of eigenvalues with respect to the design variables are not needed in the optimization process. Therefore, there is no difficulty, as shown in the examples, in finding the solutions with multiple fundamental eigenvalues.

4 Numerical Results.

Optimal topologies are found for plane trusses, and the computational efficiency and accuracy of the results are compared among the proposed method using SDP and Parametric Programming (PP) [14]. In the following example, the material of the members are steel where elastic modulus $E$ is 205.8 GPa and the mass density $\rho$ is $7.86 \times 10^{-3}$ kg/cm². In SDP, $E$ and $\rho$ are scaled so that $E = 1000.0$ is satisfied to prevent divergence in the process of finding a feasible solution. The specified eigenvalue is 1000.0 rad²/s² for all the cases. The computation has been carried out on Sun Ultra 2 Model 1300 (UltraSPARC-II 296MHz).
Table 1: Comparison of performances of SDP and PP.

| Arch \( (N^m = 174) \) \( (N^d = 106) \) | SDP \[ \text{Vol. (cm}^3 \] \( 6.4493 \times 10^5 \) \( 6.4479 \times 10^5 \) | \( \Omega_1 \) (rad\(^2\)/s\(^2\)) \( 1000.0 \) \( 982.58 \) | \( \Omega_2 \) (rad\(^2\)/s\(^2\)) \( 1000.0 \) \( 1028.2 \) | \( \bar{A}_i \) (cm\(^2\)) \( 0.0 \) \( 0.001 \) | CPU (s) \( 38.00 \) (SDPA) \( 599.01 \) (CSDP) \( 1382.81 \) |

Figure 1: A plane arch grid.

Figure 2: Optimal topology of the plane arch grid.
4.1 A plane arch grid with repeated eigenvalue

Consider a plane arch grid as shown in Fig. 1. Nonstructural masses are located at the nodes along the lowest circle. The optimal topology found by SDP after removing the members with $A_i \leq \epsilon$ is as shown in Fig. 2. The results by SDP and PP are listed in Table 1. The cross-sectional areas are linked in PP so that only symmetric designs are allowed in the optimal topology. A symmetric solution has been successfully found by SDP without assigning those additional constraints.

The multiplicity of the fundamental eigenvalues of the optimal truss is two. It should be noted here that the CPU time of PP is very large compared with that of SDP, because the number of linear equations to be solved simultaneously in PP is almost proportional to the multiplicity of the fundamental eigenvalue. We can say that the SDPA [4] is much faster than the CSDP [3] from Table 1.

5 Conclusions

Fujisawa et al. [6] reported that the SDPA is much faster than other softwares for SDPs because the SDPA incorporates data structures for handling sparse matrices and an efficient method proposed by Fujisawa et al. [5] for computing search directions when problems to be solved are large scale and sparse. Sparsity is one particular form of problem structure we exploit here because it is an important feature arising from SDP relaxation in combinatorial optimization and it has a straightforward tractable structure. However, we need an accumulation of more numerical results and knowledge for exploiting other problem structures and we consider them for future work.

The optimum design problem has been formulated as a SDP, and an algorithm has been presented for topology optimization of trusses for specified fundamental eigenvalue of vibration. The proposed algorithm is very effective for the cases of optimum designs with multiple eigenvalues, because sensitivity coefficients of the eigenvalues with respect to the design variables are not needed in the algorithm. We have seen in other examples the optimum design with up to five-fold fundamental eigenvalues can be found without any difficulty. Note that no drastic increase has been observed in CPU time for the case of multiple eigenvalues. In addition of these advantages, symmetric solutions are found without assigning any side constraints for imposing symmetricity of the cross-sectional areas.

References


