Applications of Semidefinite Optimization in Stochastic Project Scheduling

Dimitris Bertsimas, Karthik Natarajan, Chung Piaw Teo

Massachusetts Institute of Technology and National University of Singapore

Abstract— **We propose a new method, based on semidefinite optimization, to find tight upper bounds on the expected project completion time and expected project tardiness in a stochastic project scheduling environment, when only limited information in the form of first and second (joint) moments of the durations of individual activities in the project is available. Our computational experiments suggest that the bounds provided by the new method are stronger and often significant compared to the bounds found by alternative methods.**

Keywords— **Project scheduling, Problem of moments, Semidefinite programming, Co-positivity, Tardiness**

I. Introduction

A project is a set of activities that has to be completed given certain precedence relationships. We represent such a project with an acyclic directed network, in which an arc represents an activity, and a node represents the completion of all activities leading to this node. We denote the total number of arcs in the network by n and the total number of nodes by m. Node 1 represents the start of the project and node m represents the completion of the project.

For each activity (arc) i, let x_i be a random variable representing its duration, and let $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]'$. The minimum duration of the activities is assumed to be known and is represented by the vector $\boldsymbol{a} = [a_1, a_2, \ldots, a_n]'$. Clearly, $\boldsymbol{a} \geq \boldsymbol{0}$.

Let P be the set of paths in the network from node 1 to node m. For each path p let B_p be the set of activities along path p.

The project completion time is determined by the length of the longest path in the network from node 1 to node m.

$$R(\boldsymbol{x}) = \max_{p \in P} \sum_{i \in B_p} x_i.$$

A simple lower bound on the project completion time is thus

$$R(\boldsymbol{a}) = \max_{p \in P} \sum_{i \in B_p} a_i.$$

Let T be a specified due date for the project. Project tardiness is defined as

$$G(T) = (R(\boldsymbol{x}) - T)^+ = \max(0, R(\boldsymbol{x}) - T).$$

Clearly, if $T = 0$, $G(0)$ reduces to the completion time of the project. If the durations of activities are deterministic, then $G(T)$ can be exactly computed using network flow methods [9]. If the duration of activities are restricted to two possible values and are independent, Hagstrom [10] has shown that the computation of expected completion time is $\#P$-complete. The complexity status for general expected completion time is still open. However she has shown that the expected completion time cannot be computed in time polynomial in the number of values that the individual activity durations take unless $P = NP$. This suggests that in general it is difficult to compute the exact expected project tardiness hence motivating the interest in computing bounds on it.

Levy and Wiest ([16], page 169) argue that activity durations are often dependent, since different activities share the same limited resources. In practice, the information regarding the duration of the various activities is limited to knowing their expected values, and possibly their variances and covariances. Our objective in this paper is to obtain tight upper bounds on $E[(R(\boldsymbol{x}) - T)^+]$, when we only have limited information in the form of first and second (joint) moments of the durations of the various activities.

Robillard and Trahan [21] study the expected completion time $E[R(\boldsymbol{x})]$ assuming that the durations of different activities are independent random variables. Assuming independence, but only limited moment information, Devroye [7] computes upper bounds on $E[R(\boldsymbol{x})]$.

Computing lower bounds for tardiness is easier than computing upper bounds as we can use Jensen's inequality on the convex tardiness measure. This is generally observed to be reasonably accurate and tight (see Birge and Dula [4]). Computing tight upper bounds is much more challenging. One commonly used upper bound is the Edmundson-Madansky upper bound [18]. This bound uses the first moments to explicitly characterize the worst-case probability distribution. Birge and Maddox [5], building upon ideas from Mejilson and Nadas [19], develop upper bound on the expected tardiness problem when only partial information is available.

When limited moment information on the duration of various activities is available, Anklesaria [1], Sculli [22] and Mohring [17] use the central limit theorem to approximate the distribution of the tardiness. Under this method, the completion times from the various paths in the network obey a multivariate normal distribution from the central limit theorem. Then, the distribution of project tardiness can be computed by calculating the maximum of correlated normal distributions, a generally nontrivial calculation. This approach, does not provide bounds but only approximate answers. When various activities are correlated,

Research partially supported by the Singapore-MIT Alliance.

one should impose certain assumptions on the distribution of $\boldsymbol{x}$ for the central limit theorem to apply. Finally, even if the central limit theorem can be applied, it will not be a good approximation for smaller networks, often encountered in practice. In contrast, we provide formal upper bounds that are distribution free.

As the method of Birge and Maddox [5] does provide formal upper bounds, although it does not cover correlated durations of activities effectively, we briefly outline it. Assume that we know the marginal distribution functions of the duration x_i, i.e., $F_i(x) = P(x_i \leq x)$. Let Θ be the family of joint distributions compatible with the marginal distributions $F_i(x)$, $i = 1, \ldots, n$. Mejilson and Nadas [19] address the problem of evaluating

$$\sup_{\theta \in \Theta(F_1,\ldots,F_n)} E_\theta[R(\boldsymbol{x}) - T]^+. \tag{1}$$

Mejilson and Nadas [19] reformulate the problem as follows. For each path $p \in P$ and vector $\boldsymbol{z} \in \Re^n$ we have

$$\sum_{i \in B_p} x_i - T = \sum_{i \in B_p} z_i - T + \sum_{i \in B_p} [x_i - z_i].$$

Replacing $\sum_{i \in B_p} z_i - T$ by $[\max_{p \in P} \sum_{i \in B_p} z_i - T]^+$ and $\sum_{i \in B_p} [x_i - z_i]$ by $\sum_{i=1}^n [x_i - z_i]^+$ we obtain the following inequality

$$\sum_{i \in B_p} x_i - T \leq \left[\max_{p \in P} \sum_{i \in B_p} z_i - T \right]^+ + \sum_{i=1}^n [x_i - z_i]^+.$$

Since the right hand side of the inequality above is nonnegative and independent of the path p we have

$$\left[\max_{p \in P} \sum_{i \in B_p} x_i - T \right]^+ \leq \left[\max_{p \in P} \sum_{i \in B_p} z_i - T \right]^+ + \sum_{i=1}^n [x_i - z_i]^+.$$

Taking expectations and the infimum over $\boldsymbol{z} \in \Re^n$ we obtain

$$E[R(\boldsymbol{x}) - T]^+ \leq \inf_{\boldsymbol{z} \in \Re^n} \left([R(\boldsymbol{z}) - T]^+ + \sum_{i=1}^n E[x_i - z_i]^+ \right).$$

Furthermore, Mejilson and Nadas [19] construct a joint probability distribution θ in $\Theta(F_1, \ldots, F_n)$ such that the previous upper bound is tight.

Thus, an upper bound on Problem (1) can be found by

$$\sup_{\theta \in \Theta(F_1,\ldots,F_n)} \inf_{\boldsymbol{z} \in \Re^n} \left([R(\boldsymbol{z}) - T]^+ + \sum_{i=1}^n E[x_i - z_i]^+ \right) \leq$$

$$\inf_{\boldsymbol{z} \in \Re^n} \left([R(\boldsymbol{z}) - T]^+ + \sum_{i=1}^n E_{F_i}[x_i - z_i]^+ \right),$$

where the order of the sup and the inf are exchanged and it is realized that the problem inside the sup simplifies. The above problem is a nonlinear optimization problem that Birge and Maddox [5] solve using a successive linearization procedure. While in principle the previous method could be generalized to accommodate correlations among different durations, Birge and Maddox ([5], page 849) state: *"Finding efficient solutions for the problem with limited correlation information is an area for future research"*.

In this paper, building upon the work of Bertsimas and Popescu [2] and [3] we propose a different method based on semidefinite optimization to find tight upper bounds in Problem (1), when the first two moments of the durations of various activities, and the correlations are given. We develop our method in the section II and present computational results in Section III. We discuss extensions and future research in in Section IV.

II. A semidefinite formulation

Suppose that we are given $\mu_i = E[x_i]$ and $\sigma_{ij} = E[x_i x_j]$. Note that μ_i and σ_{ii} represent the first and second moments of the duration of activity i. Problem (1), when only μ_i and σ_{ij} and the lower bounds $\boldsymbol{a}$ are given, becomes

$$\begin{aligned} \max \quad & E[R(\boldsymbol{x}) - T]^+ \\ \text{s.t.} \quad & E[x_i] = \mu_i \quad 1 \leq i \leq n \\ & E[x_i x_j] = \sigma_{ij} \quad 1 \leq i \leq j \leq n \\ & \boldsymbol{x} \leq \boldsymbol{a}. \end{aligned}$$

To simplify the notation, we let $M = \binom{n+1}{2} + n$ and we define appropriate matrices $\boldsymbol{A_i}$ vectors $\boldsymbol{b_i}$ and scalars q_i such that the previous problem becomes

$$\begin{aligned} \max \quad & E[R(\boldsymbol{x}) - T]^+ \\ \text{s.t.} \quad & E[\boldsymbol{x}'\boldsymbol{A_i}\boldsymbol{x} + \boldsymbol{b_i}'\boldsymbol{x}] = q_i \quad 1 \leq i \leq M \\ & \boldsymbol{x} \geq \boldsymbol{a}. \end{aligned}$$

Simple variable transformation obtained by substituting $\boldsymbol{x} = \boldsymbol{w} + \boldsymbol{a}$ yields

$$\begin{aligned} \max \quad & E[R(\boldsymbol{w} + \boldsymbol{a}) - T]^+ \\ \text{s.t.} \quad & 1 \leq i \leq M \\ & E[\boldsymbol{w}'\boldsymbol{A_i}\boldsymbol{w} + \boldsymbol{w}'(2\boldsymbol{A_i}\boldsymbol{a} + \boldsymbol{b_i}) + \boldsymbol{a}'\boldsymbol{A_i}\boldsymbol{a} + \boldsymbol{a}'\boldsymbol{b_i}] = q_i \\ & \boldsymbol{w} \geq 0. \end{aligned}$$

Following the approach in Bertsimas and Popescu [2], [3] we introduce a dual variable y_0 that corresponds to the probability mass constraint and a dual variable y_i associated with each equality constraint and consider the dual problem:

$$\begin{aligned} \min \quad & y_0 + \sum_{i=1}^M y_i q_i \\ \text{s.t.} \quad & y_0 + \sum_{i=1}^M y_i(\boldsymbol{w}'\boldsymbol{A_i}\boldsymbol{w} + \boldsymbol{w}'(2\boldsymbol{A_i}\boldsymbol{a} + \boldsymbol{b_i}) + \boldsymbol{a}'\boldsymbol{A_i}\boldsymbol{a} + \boldsymbol{a}'\boldsymbol{b_i}) \\ & \geq (R(\boldsymbol{w} + \boldsymbol{a}) - T)^+ \quad \forall\ \boldsymbol{w} \geq \boldsymbol{0}. \end{aligned} \tag{2}$$

Haneveld [11] proves that for upper semicontinuous functions $(R(\cdot) - T)^+$ with absolute value bounded by an integrable separable function, strong duality holds. Since the tardiness function is an upper semicontinuous function, the optimal objective values of both the primal and the dual formulations are equal. Hence, the objective function value of Problem (2) is indeed $max\ E[G(T)]$.

Let path p be defined by the vector $\boldsymbol{e_p}$ where $e_p^i = 1$ if activity i is on path p, and 0, otherwise. Let $\boldsymbol{e_0}$ be a zero vector in $\Re^n$. We also define $d_p = -T, \forall\, p \in P$ and $d_0 = 0$. With these definitions the computation of the upper bound on the expected tardiness can we can be rewritten as:

$$\begin{aligned} \min \quad & y_0 + \sum_{i=1}^{M} y_i q_i \\ \text{s.t.} \quad & y_0 + \sum_{i=1}^{M} y_i(\boldsymbol{w'A_i w} + \boldsymbol{w}'(2\boldsymbol{A_i a} + \boldsymbol{b_i}) + \boldsymbol{a'A_i a} + \boldsymbol{a'b_i}) \\ & \geq \boldsymbol{e_p}'(\boldsymbol{w} + \boldsymbol{a}) + d_p \ \forall \boldsymbol{w} \in \Re^n_+, p \in P \cup \{0\}. \end{aligned} \qquad (3)$$

We define

$$\begin{aligned} L_1(\boldsymbol{y}) &= \sum_{i=1}^{M} y_i \boldsymbol{A_i}, \\ L_{2p}(\boldsymbol{y}) &= \left(\sum_{i=1}^{M} y_i(2\boldsymbol{A_i a} + \boldsymbol{b_i}) - \boldsymbol{e_p} \right) /2, \\ L_{3p}(\boldsymbol{y}) &= y_0 + \sum_{i=1}^{M} y_i(\boldsymbol{a'A_i a} + \boldsymbol{a'b_i}) - \boldsymbol{e_p}'\boldsymbol{a} - d_p. \end{aligned}$$

Here L_1 is a symmetric matrix, and is independent of p, while $L_{2p} \in \Re^n$ and $L_{3p} \in \Re$ depend on p. With these definitions $max\ E[G(T)]$ is equal to:

$$\begin{aligned} \min \quad & y_0 + \sum_{i=1}^{M} y_i q_i \\ \text{s.t.} \quad & \forall \boldsymbol{w} \in \Re^n_+, \ \forall p \in P \cup \{0\} \\ & \begin{pmatrix} \boldsymbol{w} \\ 1 \end{pmatrix}' \begin{pmatrix} L_1(\boldsymbol{y}) & L_{2p}(\boldsymbol{y}) \\ L_{2p}(\boldsymbol{y})' & L_{3p}(\boldsymbol{y}) \end{pmatrix} \begin{pmatrix} \boldsymbol{w} \\ 1 \end{pmatrix} \geq 0. \end{aligned} \qquad (4)$$

This implies that the matrix in (4) belongs in the cone of co-positive matrices $C^+_{n+1} = \{\boldsymbol{A} |\ \boldsymbol{y'Ay} \geq 0, \ \forall\ \boldsymbol{y} \in \Re^n_+\}$. Hence $max\ E[G(T)]$ is equal to:

$$\begin{aligned} \min \quad & y_0 + \sum_{i=1}^{M} y_i q_i \\ \text{s.t.} \quad & \begin{pmatrix} L_1(\boldsymbol{y}) & L_{2p}(\boldsymbol{y}) \\ L_{2p}(\boldsymbol{y})' & L_{3p}(\boldsymbol{y}) \end{pmatrix} \in C^+_{n+1} \ \forall p \in P \cup \{0\}. \end{aligned} \qquad (5)$$

Determining if a given matrix is co-positive is $co-NP$-complete (see Kabadi and Murty [13]) . Thus, an exact tractable description of the co-positive cone is not known, and most probably impossible unless $P = co-NP$. For this reason, we find sufficient conditions for co-positivity. A simple such condition is as follows. If $\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{C}$, such that the matrix $\boldsymbol{B}$ is positive semidefinite, denoted by $\boldsymbol{B} \succeq \boldsymbol{0}$, and the matrix $\boldsymbol{C}$ has non-negative entries, denoted by $\boldsymbol{C} \geq \boldsymbol{0}$, then clearly the matrix $\boldsymbol{A}$ is co-positive. Thus, the following semidefinite optimization problem gives an upper bound on the expected tardiness $G(T)$.

$$\begin{aligned} \min \quad & y_0 + \sum_{i=1}^{M} y_i q_i \\ \text{s.t.} \quad & \boldsymbol{N_p} \geq \boldsymbol{0} \ \ \forall p \in P \cup \{0\} \\ & \boldsymbol{N_p} = \boldsymbol{N_p}' \ \ \forall p \in P \cup \{0\} \\ & \begin{pmatrix} L_1(\boldsymbol{y}) & L_{2p}(\boldsymbol{y}) \\ L_{2p}(\boldsymbol{y})' & L_{3p}(\boldsymbol{y}) \end{pmatrix} \succeq \boldsymbol{N_p} \ \ \forall p \in P \cup \{0\}. \end{aligned} \qquad (6)$$

Note that the total number of semidefinite constraints is equal to the total number of paths in the network plus one, $|P| + 1$.

III. Computational Results

In this section, we compare the bounds provided by the semidefinite model (6) and by the method in Birge and Maddox [5] on two small examples in [5]. The semidefinite problem (6) is solved using the software Sedumi developed by Sturm [23].

A. Example 1

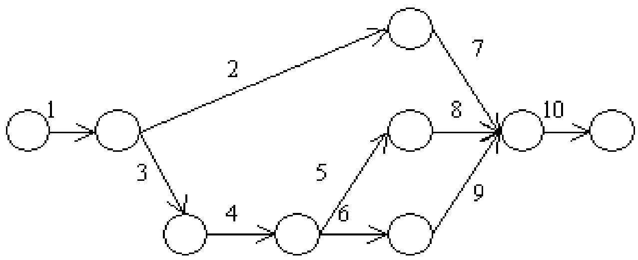

Figure 1. The project network in Example 1.

The project above consists of 10 activities that need to be completed and consists of 3 paths. The given first and second moments with the lower bound on the arc durations is provided in Table I. The upper bounds for five different due dates were computed for the network above and compared with the results reported in Birge and Maddox [5]. The results obtained are displayed in Table II.

TABLE I
Activity duration data for Example 1

Activity i	a_i	$E[x_i]$	$E[x_i]^2$
1	1	1	1
2	2	3	9.333
3	1	2	4.333
4	2	2.5	6.333
5	3	5	26.333
6	3	4	16.333
7	1	3	10.333
8	4	4.5	20.333
9	1	1.5	2.333
10	4	5	25.333

TABLE II

Upper bounds on project tardiness for Example 1.

Due Date T	0	15	18.33	21.67	25
BM [5]	20.35	5.35	2.98	1.27	0.73
SDP	20.23	5.30	2.67	0.95	0.48

B. *Example 2*

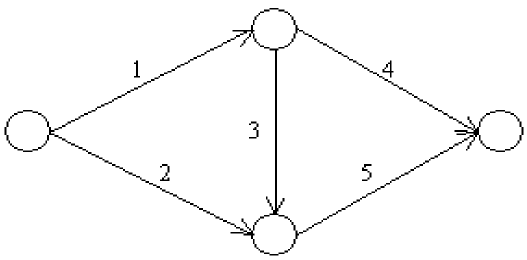

Figure 2. The project network in Example 2.

This project is a smaller project as compared to Example 1 with 5 activities. The given data for this project has been provided in the Table III. We evaluate the project tardiness

TABLE III

Activity duration data for Example 2

Activity i	a_i	$E[x_i]$	$E[x_i]^2$
1	0	1	1.666
2	0	1	1.666
3	0	1	1.666
4	0	1	1.666
5	0	1	1.666

for four deadline date values in Table IV.

TABLE IV

Upper bounds on project tardiness for Example 2.

Due Date T	0	2	4	6
BM [5]	5	2.27	1.16	0.73
SDP	4.16	2.25	1.01	0.55

As evident from Table II and IV the bound provided by solving Problem (6), always outperforms the bound obtained by Birge and Maddox [5]. Thus, even for problems with known first and second moments only, we obtain stronger upper bounds on project tardiness as compared to the technique described in Section II. This makes the technique promising as it is capable of incorporating cross moment information as well.

IV. Extensions and future research

The approach outlined in the paper can be used to handle several extensions:

(a) If higher moment information is available, then analogously to Eq.3) the dual problem can be written as a multivariate polynomial $P(\boldsymbol{w})$, whose coefficients are linear functions of the dual variables $\boldsymbol{y}$ is nonnegative for all $\boldsymbol{w} \in \Re^n_+$. A sufficient, but not necessary, condition that a multivariate polynomial is nonnegative is that can be expressed as a sum of squares of polynomials. This leads to a semidefinite formulation, see Parillo [20]. Hence we can find the worst-case bounds on project tardiness if higher moment information is known.

(b) If we want to penalize the delay of a project after its deadline date more severely, then we could choose a piecewise quadratic function as the tardiness instead of a piecewise linear function. Such definitions of project tardiness can be incorporated in the proposed approach and can be handled efficiently.

(c) Finding an upper bound on the variance of the project completion time or on the probability of a project being overdue, namely $P(R(\boldsymbol{x}) > T)$, is also possible under our approach.

The principle limitation of the current technique is that the number of semidefinite constraints it generates is proportional to the number of paths. This path dependent formulation may be too huge for large networks. Though the state of art semidefinite programming software is relatively advanced we need techniques to effectively handle this path dependency in the formulation. One such idea may be to evaluate heuristics to determine paths that are most likely of becoming the critical path. This could be then used to reduce the number of semidefinite constraints [1], [22]. Another promising technique could be one that uses cutting plane techniques to solve the semidefinite program with huge number of semidefinite constraints effectively [14]. It is necessary to evaluate and develop these techniques and then test them effectively on large projects with large number of paths.

References

[1] K.P. Anklesaria and Z. Drezner. "A multivariate approach to estimating the completion time for PERT networks", *Journal of the Operational Research Society* 40: 811-815, 1986.

[2] D. Bertsimas and I. Popescu. "Optimal inequalities in probability: A convex optimization approach", 1999.

[3] D. Bertsimas and I. Popescu. "On the relation between option and stock prices: A convex optimization approach", 1999.

[4] J.R. Birge and J.H. Dula. "Bounding separable recourse functions with limited distribution information", *Annals of Operations Research* 30: 277-298, 1991.

[5] J.R. Birge and M.J. Maddox. "Bounds on expected project tardiness", *Operations Research* 43: 838-850, 1995.

[6] I.M. Bomze and E.d. Clerk. "Solving standard quadratic optimization problems via linear, semidefinite and co-positive programming", 2001.

[7] L.P. Devroye. "Inequalities for the completion times of stochastic PERT networks", *Mathematics of Operations Research* 4 No 4: 441-447, 1979.

[8] J.H. Dula and R.V. Murthy. "A Tchebysheff-type bound on the expectation of sublinear polyhedral functions", *Operations Research* 40 No 5: 914-922, 1992.

[9] S.E. Elmaghraby. "Project planning and control by network models", John Wiley and Sons, 1977.

[10] J.H. Hagstrom. "Computational complexity of PERT problems", *Networks* 18: 139-147, 1988.

[11] W.K.K. Haneveld. "Robustness against dependence in PERT: An application of duality and distributions with known marginals", *Mathematical Programming Study* 27: 153-182, 1986.
[12] K. Isii. "On the sharpness of chebyshev-type inequalities", *Ann. Inst. Stat. Math* 14: 185-197, 1963.
[13] S.N. Kabadi and K.G. Murty. "Some NP-complete problems in quadratic and nonlinear programming", *Mathematical Programming* 39: 117-129, 1987.
[14] K. Krishnan and J.E. Mitchell. "A linear programming approach to semidefinite programs", 2001.
[15] J.B. Lasserre. "Global optimization with polynomials and the problem of moments", *SIAM Journal of Optimization* 11: 796-817, 2001.
[16] F.K. Levy and J.D. Wiest. "A Management guide to PERT/CPM: with GERT/PDM/DPCM and other networks", Prentice Hall 2nd ed, 1977.
[17] A. Ludwig, R.H. Mohring and F. Stork. "A computational study on bounding the makespan distribution in stochastic project networks", 1998.
[18] A. Madansky. "Inequalities for stochastic linear programming problems", *Management Science* 6: 197-204, 1960
[19] I. Mejilson and A. Nadas. "Convex majorization with an application to the length of critical path", *Journal of Applied Probability* 16: 671-677, 1979.
[20] P.A. Parillo. "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization", PhD Thesis, California Institute of Technology 2000
[21] P. Robillard and M. Trahan. "The completion times of PERT networks", *Operations Research* 25: 15-29, 1976.
[22] D. Sculli and Y.W. Shum. "An approximate solution to the PERT problem", *Computers and Mathematics with Application* 21: 1-7, 1991.
[23] J.F. Sturm. SeDuMi version 1.03, Available from http://fewcal.kub.nl/sturm/software/sedumi.html.
[24] G. Weiss. "Stochastic bounds on distributions of optimal value functions with applications to PERT, network flows and reliability", *Operations Research* 34 No 4: 595-605, 1986.