2160/2004

Raport Badawczy

Research Report

RB/66/2004

On Linear Time Algorithms for the Continuous Quadratic Knapsack Problem

Krzysztof C. Kiwiel

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 8373578

fax: (+48) (22) 8372772

Kierownik Pracowni zgłaszający pracę:
Prof. dr hab. inż. Krzysztof C. Kiwiel

Warszawa 2004

On linear time algorithms for the continuous quadratic knapsack problem*

Krzysztof C. Kiwiel†

March 27, 2004

Abstract

We give a linear time algorithm for the continuous quadratic knapsack problem which is both simpler than the existing methods and competitive in practice. Encouraging computational results are presented for large-scale problems.

Key words. Nonlinear programming, convex programming, quadratic programming, separable programming, singly constrained quadratic program.

1 Introduction

The *continuous quadratic knapsack problem* is defined by

$$\mathrm{P}: \quad \min \quad f(x) := \tfrac{1}{2}x^T Dx - a^T x \quad \text{s.t.} \quad b^T x = r, \quad l \le x \le u, \tag{1.1}$$

where x is an n-vector of variables, $a, b, l, u \in \mathbb{R}^n$, $r \in \mathbb{R}$, $D = \operatorname{diag}(d)$ with $d > 0$, so that the objective f is strictly convex. Assuming P is feasible, let x^* denote its unique solution.

Problem P has many applications; see, e.g., [Bru84, BSS95, CaM87, MSMJ03, PaK90] and references therein.

Specialized algorithms for P solve its dual problem by finding a Lagrange multiplier t_* that solves the equation $g(t) = r$, where g is a monotone piecewise linear function with $2n$ breakpoints (cf. §2). To this end, the $O(n)$ algorithms of [Bru84, CaM87, PaK90, MSMJ03] use medians of breakpoint subsets. However, they are quite complicated, and the analysis of [MSMJ03, PaK90] has some gaps that are not easy to fix (cf. Rem. 2.3(d)).

In this paper we introduce a simpler $O(n)$ algorithm that is both easier to analyze and competitive in practice with those in [Bru84, CaM87, MSMJ03, PaK90].

The paper is organized as follows. In §2 we review some properties of P and present our method. Additional constructions of [CaM87] are discussed in §3. Finally, computational results for large-scale problems are reported in §4.

*Research supported by the State Committee for Scientific Research under Grant 8T11A00115.

†Systems Research Institute, Newelska 6, 01–447 Warsaw, Poland (kiwiel@ibspan.waw.pl)

2 The breakpoint searching algorithm

Viewing $t \in \mathbb{R}$ as a multiplier for the equality constraint of P in (1.1), consider the *Lagrangian primal solution* (the minimizer of $f(x) + t(b^T x - r)$ s.t. $l \le x \le u$)

$$x(t) := \min\left\{ \max\left[l, D^{-1}(a - tb) \right], u \right\} \tag{2.1}$$

(where the min and max are taken componentwise) and its *constraint value*

$$g(t) := b^T x(t). \tag{2.2}$$

Solving P amounts to solving $g(t) = r$. Indeed, invoking the Karush-Kuhn-Tucker conditions for P as in [CaM87, Thm 2.1], [PaK90, Thm 2.1] gives the following result.

Fact 2.1. *$x^* = x(t)$ iff $g(t) = r$. Further, the set $T_* := \{\, t : g(t) = r \,\}$ is nonempty.*

As in [Bru84], we assume for simplicity that $b > 0$, because if $b_i = 0$, x_i may be eliminated ($x_i^* = \min\{\max[l_i, a_i/d_i], u_i\}$), whereas if $b_i < 0$, we may replace $\{x_i, a_i, b_i, l_i, u_i\}$ by $-\{x_i, a_i, b_i, u_i, l_i\}$ (in fact this transformation may be *implicit*).

By (2.1)-(2.2), the function g has the following *breakpoints*

$$t_i^l := (a_i - l_i d_i)/b_i \quad \text{and} \quad t_i^u := (a_i - u_i d_i)/b_i, \quad i = 1{:}\,n, \tag{2.3}$$

with $t_i^u \le t_i^l$ (from $l_i \le u_i$ and $b_i > 0$), and each $x_i(t)$ may be expressed as

$$x_i(t) = \begin{cases} u_i & \text{if } t \le t_i^u, \\ (a_i - tb_i)/d_i & \text{if } t_i^u \le t \le t_i^l, \\ l_i & \text{if } t_i^l \le t. \end{cases} \tag{2.4}$$

Thus $g(t)$ is a continuous, piecewise linear and *nonincreasing* function of t.

To locate an optimal t_* in T_*, the algorithm below generates a *bracketing interval* $[t_L, t_U]$ that contains T_* by evaluating g at median breakpoints in (t_L, t_U) until (t_L, t_U) contains no breakpoints; then g is linear on $[t_L, t_U]$, and t_* is found by interpolation.

Algorithm 2.2.

Step 0 (*Initiation*). Set $N := \{1{:}\,n\}$, $T := \{t_i^l\}_{i\in N} \cup \{t_i^u\}_{i\in N}$, $t_L := -\infty$, $t_U := \infty$.

Step 1 (*Breakpoint selection*). Set $\hat{t} :=$ median(T) (the median of the set T).

Step 2 (*Computing $g(\hat{t})$*). Calculate $g(\hat{t})$.

Step 3 (*Optimality check*). If $g(\hat{t}) = r$, stop with $t_* := \hat{t}$.

Step 4 (*Lower breakpoint removal*). If $g(\hat{t}) > r$, set $t_L := \hat{t}$, $T := \{t \in T : \hat{t} < t\}$.

Step 5 (*Upper breakpoint removal*). If $g(\hat{t}) < r$, set $t_U := \hat{t}$, $T := \{t \in T : t < \hat{t}\}$.

Step 6 (*Stopping criterion*). If $T \ne \emptyset$, go to Step 1; otherwise, stop with

$$t_* := t_L - [\,g(t_L) - r\,]\frac{t_U - t_L}{g(t_U) - g(t_L)}. \tag{2.5}$$

The following comments clarify the nature of the algorithm.

Remarks 2.3. (a) By the argument of [CaM87, p. 1438], Algorithm 2.2 only requires order n operations, since $|T|$ is originally $2n$, $\hat{t} :=$ median(T) can be obtained in order $|T|$ operations [Knu98, §5.3.3], the evaluation of $g(\hat{t})$ requires order $|T|$ operations (see below), and each iteration reduces $|T|$ at least by half at Steps 4 or 5.

(b) To compute $g(\hat{t})$ efficiently, we may partition the set N into the following sets

$$L := \left\{ i : t_i^l \le t_L \right\}, \quad M := \left\{ i : t_L, t_U \in \left[t_i^u, t_i^l \right] \right\}, \quad U := \left\{ i : t_U \le t_i^u \right\}, \tag{2.6a}$$

$$I := \left\{ i : t_i^l \in (t_L, t_U) \text{ or } t_i^u \in (t_L, t_U) \right\}; \tag{2.6b}$$

note that $|I| \le |T| \le 2|I|$. Thus, by (2.2), (2.4) and (2.6),

$$g(t) = \sum_{i \in I} b_i x_i(t) + (p - tq) + s \quad \forall t \in [t_L, t_U], \tag{2.7}$$

where

$$\sum_{i \in I} b_i x_i(t) = \sum_{i \in I : t \in [t_i^u, t_i^l]} b_i(a_i - tb_i)/d_i + \sum_{i \in I : t_i^l < t} b_i l_i + \sum_{i \in I : t < t_i^u} b_i u_i, \tag{2.8}$$

$$p := \sum_{i \in M} a_i b_i / d_i, \quad q := \sum_{i \in M} b_i^2 / d_i \quad \text{and} \quad s := \sum_{i \in L} b_i l_i + \sum_{i \in U} b_i u_i. \tag{2.9}$$

Setting $I := N$, $p, q, s := 0$ at Step 0, at Step 6 we may update I, p, q and s as follows:

```
for $i \in I$ do
    if $t_i^l \le t_L$, set $I := I \setminus \{i\}$, $s := s + b_i l_i$;
    if $t_U \le t_i^u$, set $I := I \setminus \{i\}$, $s := s + b_i u_i$;                (2.10)
    if $t_L, t_U \in [t_i^u, t_i^l]$, set $I := I \setminus \{i\}$, $p := p + a_i b_i / d_i$, $q := q + b_i^2 / d_i$.
```

This update and the calculation of $g(\hat{t})$ require order $|I| \le |T|$ operations.

(c) Upon termination, $x^* = x(t_*)$ is recovered via (2.1) in order n operations.

(d) The algorithm of [PaK90] is quite similar to ours, but it fails on simple examples (e.g., for $n = 2$, $d = b = (1, 1)$, $a = 0$, $r = -2$, $l = (-2, -2)$, $u = (-1, 0)$). The algorithm of [MSMJ03] is much more complicated, and may also fail (e.g., on the example of [MSMJ03, pp. 565–566]).

3 Further breakpoint removal of Calamai and Moré

The original version of [CaM87, Alg. 2.3] corresponds to replacing Steps 4 and 5 by

Step 4' (*Lower breakpoint removal*). If $g(\hat{t}) > r$ then find the right adjacent breakpoint $\check{t} := \min\{t \in T : \hat{t} < t\}$; if $\check{t} < \infty$ and $g(\check{t}) > r$, set $t_L := \check{t}$, $T := \{t \in T : \check{t} \le t\}$, else set $t_L := \hat{t}$, $t_U := \min\{t_U, \check{t}\}$ and stop with t_* given by (2.5).

Step 5' (*Upper breakpoint removal*). If $g(\hat{t}) < r$ then find the left adjacent breakpoint $\check{t} := \max\{t \in T : t < \hat{t}\}$; if $\check{t} > -\infty$ and $g(\check{t}) < r$, set $t_U := \check{t}$, $T := \{t \in T : t \le \check{t}\}$, else set $t_L := \max\{t_L, \check{t}\}$, $t_U := \hat{t}$ and stop with t_* given by (2.5).

Table 4.1: Run times of Algorithm 2.2 (in seconds).

	uncorrelated			weakly correl.			strongly correl.			overall		
n	avg	max	min	avg	max	min	avg	max	min	avg	max	min
50000	0.04	0.05	0.04	0.04	0.05	0.03	0.04	0.05	0.04	0.04	0.05	0.03
100000	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.08	0.07	0.08	0.08	0.07
500000	0.36	0.41	0.33	0.36	0.38	0.35	0.37	0.38	0.35	0.36	0.41	0.33
1000000	0.72	0.74	0.65	0.71	0.74	0.67	0.72	0.74	0.67	0.72	0.74	0.65
1500000	1.06	1.10	0.98	1.06	1.10	1.00	1.06	1.10	0.99	1.06	1.10	0.98
2000000	1.42	1.46	1.31	1.41	1.46	1.29	1.42	1.46	1.30	1.42	1.46	1.29

By (2.4) and (2.7), because $\hat{t}$ and $\check{t}$ are *consecutive* breakpoints, we may compute

$$g(\check{t}) = g(\hat{t}) - (\check{t} - \hat{t})[q + \check{q}] \quad \text{with} \quad \check{q} := \sum_{i \in I : \hat{t}, \check{t} \in [t_i^u, t_i^l]} b_i^2 / d_i \tag{3.1}$$

in order $|I|$ operations. Yet this modification will typically remove only one more breakpoint, and this may not be worth the additional effort in finding $\check{t}$ and $g(\check{t})$.

4 Numerical results

Algorithm 2.2, the Calamai-Moré version of §3 and Brucker's method [Bru84] were programmed in Fortran 77 and run on a notebook PC (Pentium 4M 2 GHz, 768 MB RAM) under MS Windows XP. We used the median finding routine of [Kiw03], which permutes T to place elements $< \hat{t}$ first, then elements $= \hat{t}$, and finally elements $> \hat{t}$. Hence the updates of Steps 4 and 5 only require a change of one pointer.

Our test problems were randomly generated with n ranging between 50000 and 2000000. As in [BSS95, §2], all parameters were distributed uniformly in the intervals of the following three problem classes: (1) uncorrelated: $a_i, b_i, d_i \in [10, 25]$; (2) weakly correlated: $b_i \in [10, 25]$, $a_i, d_i \in [b_i - 5, b_i + 5]$; (3) strongly correlated: $b_i \in [10, 25]$, $a_i = d_i = b_i + 5$; further, $l_i, u_i \in [1, 15]$, $i \in N$, $r \in [b^T l, b^T u]$. For each problem size, 20 instances were generated in each class.

Tables 4.1–4.3 report the average, maximum and minimum run times over the 20 instances for each of the listed problem sizes and classes, as well as overall statistics. The average run times grow linearly with the problem size. The Calmai-Moré algorithm and Brucker's one were slower than our method by about 30% and 16%, respectively.

References

[Bru84] P. Brucker, *An $O(n)$ algorithm for quadratic knapsack problems*, Oper. Res. Lett. **3** (1984) 163–166.

[BSS95] K. M. Bretthauer, B. Shetty and S. Syam, *A branch and bound algorithm for integer quadratic knapsack problems*, ORSA J. Comput. **7** (1995) 109–116.

Table 4.2: Run times of the Calamai-Moré algorithm (in seconds).

	uncorrelated			weakly correl.			strongly correl.			overall		
n	avg	max	min	avg	max	min	avg	max	min	avg	max	min
50000	0.05	0.07	0.05	0.05	0.06	0.05	0.05	0.06	0.05	0.05	0.07	0.05
100000	0.10	0.11	0.09	0.10	0.10	0.09	0.10	0.10	0.09	0.10	0.11	0.09
500000	0.47	0.51	0.45	0.47	0.48	0.46	0.48	0.49	0.46	0.47	0.51	0.45
1000000	0.94	0.97	0.88	0.94	0.96	0.89	0.94	0.96	0.92	0.94	0.97	0.88
1500000	1.40	1.44	1.32	1.40	1.43	1.35	1.39	1.44	1.33	1.40	1.44	1.32
2000000	1.86	1.90	1.75	1.86	1.90	1.74	1.86	1.89	1.75	1.86	1.90	1.74

Table 4.3: Run times of Brucker's algorithm (in seconds).

	uncorrelated			weakly correl.			strongly correl.			overall		
n	avg	max	min	avg	max	min	avg	max	min	avg	max	min
50000	0.05	0.07	0.03	0.05	0.05	0.04	0.05	0.05	0.04	0.05	0.07	0.03
100000	0.09	0.10	0.07	0.09	0.10	0.07	0.09	0.10	0.07	0.09	0.10	0.07
500000	0.42	0.47	0.30	0.43	0.47	0.36	0.42	0.45	0.34	0.42	0.47	0.30
1000000	0.82	0.93	0.61	0.85	1.00	0.65	0.81	0.91	0.64	0.83	1.00	0.61
1500000	1.25	1.36	1.06	1.26	1.49	1.07	1.23	1.35	0.99	1.25	1.49	0.99
2000000	1.67	1.82	1.37	1.62	1.85	1.25	1.66	1.83	1.36	1.65	1.85	1.25

[CaM87] P. H. Calamai and J. J. Moré, *Quasi-Newton updates with bounds*, SIAM J. Numer. Anal. **24** (1987) 1434–1441.

[Kiw03] K. C. Kiwiel, *Randomized selection with quintary partitions*, Tech. report, Systems Research Institute, Warsaw, 2003. Available at the URL http://arxiv.org/abs/cs.DS/0312055.

[Knu98] D. E. Knuth, *The Art of Computer Programming. Volume III: Sorting and Searching*, second ed., Addison-Wesley, Reading, MA, 1998.

[MSMJ03] N. Maculan, C. P. Santiago, E. M. Macambira and M. H. C. Jardim, *An $O(n)$ algorithm for projecting a vector on the intersection of a hyperplane and a box in R^n*, J. Optim. Theory Appl. **117** (2003) 553–574.

[PaK90] P. M. Pardalos and N. Kovoor, *An algorithm for a singly constrained class of quadratic programs subject to upper and lower bounds*, Math. Programming **46** (1990) 321–328.