On the Number of Pivots of Dantzig's Simplex Methods for Linear and Convex Quadratic Programs

Xinyao Zhang

Joint work with Shaoning Han and Jong-Shi Pang

University of Southern California

2024 INFORMS Optimization Society Conference

March 2024

Agenda

Background and Motivation

- **Goal:**
  - **Find upper-bounds for the number of simplex method pivots**
  - **based on quantities derived from problem inputs**
- Different from
  - the worst-case exponential complexity
  - the probabilistic average-case analysis
  - the polynomial complexity of the ellipsoid method or interior-point approaches bounded by the input size of the problem data
  - ...
- Inspired by the work of Ye (2011) on Markov decision problems.
- Expansion of Kitahara-Mizuno's analysis for linear programs (2012).

MDP result

Markov Decision Problem with a fixed discount rate $0 \leq \theta < 1$:

$$\begin{aligned} \min \quad & c^T x, \\ \text{subject to} \quad & [I - \theta P_1, I - \theta P_2, ..., I - \theta P_k] x = e, \\ & x \geq 0. \end{aligned}$$

e: the vector of all ones. k: the number of possible actions.
P_i: a $m \times m$ Markov matrix, i.e. $e^T P_i = e^T$.
$A \in R^{m \times km}$, $b = e \in R^m$, $c \in R^{km}$, $x \in R^{km}$.

MDP result [Ye, 2011]

The simplex method with the most-negative rule for solving the MDP with fixed discount rate $0 \leq \theta < 1$ terminates in at most $\frac{m^2(k-1)}{1-\theta} log(\frac{m^2}{1-\theta})$ iterations.

General LP result

Linear Program (LP)

$$\text{minimize } c^T x$$

$$\text{subject to } Ax = b \text{ and } x \geq 0, \text{ where } A \in \mathbb{R}^{m \times n},$$

General LP result [Kitahara et al., 2012]

An upper bound for the number of different BFSs generated by the simplex method with the most-negative rule or the best-improvement rule:

$$n \lceil m \frac{\gamma}{\delta} \log(m \frac{\gamma}{\delta}) \rceil$$

δ: the minimum value of all the positive elements of primal BFSs,
γ: the maximum value of all the positive elements of primal BFSs.

* Primal nondegeneracy $\Rightarrow$ an upper bound for the number of iterations.

Our Contributions

- Specialize the KM+ bound to Leontief systems and to an LCP with a hidden Z-matrix.
- Derive iteration bounds for the Dantzig (1961) and van de Panne-Whinston (1964) (pivoting) algorithm for nonnegatively convex quadratic programs:

$$\text{minimize } \frac{1}{2}z^T Mz + q^T z, \quad z \in \mathbb{R}^n_+.$$

- Provide two classes of QPs to illustrate the derived bounds

Leontief Systems

The Leontief systems and associated linear programs,

$$
\begin{aligned}
\min\ & c^\top x \\
\text{s.t.}\ & Ax = b \\
& x \geq 0,
\end{aligned} \tag{1}
$$

where $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are given vectors and the matrix $A \in \mathbb{R}^{m \times n}$ satisfies the following conditions:

(a) Each column of A has at most one positive entry.

(b) $b \in \mathbb{R}^m_{++}$.

(c) A feasible solution of (1) exists.

A is *pre-Leontief-plus* when each column has exactly one positive entry.

Leontief Systems

Pre-Leontief plus constraint matrix

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
$+$	$+$	$+$	$+$	$\ominus$	$\ominus$	$\ominus$	$\ominus$	$\ominus$	$\ominus$
$\ominus$	$\ominus$	$\ominus$	$\ominus$	$+$	$+$	$\ominus$	$\ominus$	$\ominus$	$\ominus$
$\ominus$	$\ominus$	$\ominus$	$\ominus$	$\ominus$	$\ominus$	$+$	$+$	$+$	$+$

Note: A pre-Leontief plus matrix A can be arranged to specific patterns with '$+$' for positive and '$\ominus$' for nonpositive elements.
With A structured as displayed, and assuming non-empty index groups G_g for all $g = 1, \ldots, m$, we define an $m \times m$ matrix $\bar{A}$ with entries:

$$\bar{a}_{ig} = \begin{cases} \min\limits_{j \in G_g} a_{ij} & \text{if } i = g \\ -\max\limits_{j \in G_g} |a_{ij}| & \text{if } i \neq g. \end{cases}$$

The off-diagonal entries of $\bar{A}$ are clearly nonpositive, making $\bar{A}$ a Z-matrix.
*We assume $\bar{A}$ is Minkowski for following results

Nondegeneracy of Leontief Systems

- A square matrix with nonpositive off-diagonal entries is called a Z-matrix.
- If a Z-matrix further has a nonnegative inverse, it is termed a Minkowski matrix.
- **Key Fact:** A Z-matrix M is Minkowski $\Leftrightarrow$ there exists a positive vector d such that $Mx = d$ has a nonnegative solution.

If (1) is feasible and A is *pre-Leontief-plus*, then any feasible basis B of such a system with a positive right-hand side b must be:

- A Minkowski matrix.
- **Nondegenerate**, as $B^{-1}b > 0$.

Leontief Systems

Leontief Systems result

If $A \in \mathbb{R}^{m \times n}$ is pre-Leontief-plus and the matrix $\bar{A}$ is Minkowski, then:

- if $\bar{x}$ is any basic feasible solution of the system $Ax = b, x \geq 0$, where $b \in \mathbb{R}^m_{++}$, then:

$$\underline{\delta} \triangleq \min_{1 \leq i \leq m} \left[\frac{b_i}{\max\limits_{j \in \mathcal{G}_i} a_{ij}} \right] \leq \delta_{\bar{x}} \leq \gamma_{\bar{x}} \leq \max_{1 \leq i \leq m} \left[\left(\overline{A} \right)^{-1} b \right]_i \triangleq \overline{\gamma},$$

  where δ and γ are the smallest and largest positive elements of $\bar{x}$.

- Let $C \triangleq \dfrac{\overline{\gamma}}{\underline{\delta}}$. If (1) has an optimal solution, the simplex method with the least reduced cost rule will solve (1) in no more than $n \lceil m\, C \log(m\, C) \rceil$ pivots.

Nonnegatively Convex Quadratic Programs

Convex quadratic program (QP)

$$\min_{z \in \mathbb{R}^n_+} v(z) \triangleq \frac{1}{2} z^\top M z + q^\top z, \tag{2}$$

where the matrix M is symmetric positive semidefinite and q is arbitrary.

This QP is the dual of the strictly convex QP:

$$\min_{x \in \mathbb{R}^m} \frac{1}{2} x^\top Q x + p^\top x \quad \text{subject to} \quad Ax \leq b,$$

with a positive definite $Q \in \mathbb{R}^{m \times m}$ via $M = AQ^{-1}A^\top$, $q = b + AQ^{-1}p$.
The KKT conditions for (1) are given by the LCP (q, M):

$$0 \leq z \perp w = q + Mz \geq 0,$$

* Assume finite optimal solution exists.

The DvPW Algorithm

The pivoting method for QP due independently to Dantzig [Dantzig, 1963] and to van de Panne-Whinston [Van de Panne and Whinston, 1964];

$$0 \le z \perp w = q + Mz \ge 0,$$

Step 0. Initialization

Input (q, M) with M being symmetric and positive semi-definite. Set $(q^0, M^0) = (q, M)$, $\nu = 0$, $\alpha = \emptyset$, and $\beta = \{1, \ldots, n\}$.

Step 1. Test for termination

Breaking ties arbitrarily, choose $r \in \arg\min_{i \in \beta} \{q_i^\nu\}$.

1A. If $q_r^\nu \ge 0$, stop. A solution is given by $\hat{z}$ where $z_\alpha = q_\alpha^\nu$, $z_\beta = 0$.

1B. If $q_r^\nu < 0$, let w_r^ν be the distinguished and z_r^ν be the driving variable.

1C. If $m_{rr}^\nu = 0$ and $m_{ir}^\nu \ge 0$ for all $i \in \alpha$, stop. There is no solution.

The DvPW Algorithm

Step 2. Determination of the blocking variable

Use the minimum ratio test to define the index s of a blocking variable.

While increasing z_r^ν, the following may happen

- w_r increasing and reaches 0 $\Rightarrow$ blocked by complementarity
  - Good! r fixed. Next major cycle.
- $z_{s \in \alpha}$ decreasing and reaches 0 $\Rightarrow$ blocked by primal feasibility
  - Hmm..r is not fixed yet. Lock $z_s = 0$ and unlock $w_s \uparrow$. Keep driving z_r.

Step 3. Pivoting.

The driving variable z_r^ν is blocked by w_s^ν. Pivot (w_s^ν, z_s^ν) to update $(q^{\nu+1}, M^{\nu+1})$.
If $s = r$, transfer r from β to α. Go to Step 1 with $\nu \leftarrow \nu + 1$.
If $s \neq r$, transfer s from α to β. Go to Step 2 with $\nu \leftarrow \nu + 1$.

[Cottle et al., 2009]

The DvPW Algorithm for solving the LCP (q, M)

	1	z
$q^\top z$	0	$q^\top$
w	q	M

(3)

After each principal pivot, we update the following tableau:

	1	w_α	z_β
$q^\top z$	$-q_\alpha^\top (M_{\alpha\alpha})^{-1} q_\alpha$	$q_\alpha^\top (M_{\alpha\alpha})^{-1}$	$q_\beta^\top - q_\alpha^\top (M_{\alpha\alpha})^{-1} M_{\alpha\beta}$
z_α	$-(M_{\alpha\alpha})^{-1} q_\alpha$	$(M_{\alpha\alpha})^{-1}$	$-(M_{\alpha\alpha})^{-1} M_{\alpha\beta}$
w_β	$q_\beta - M_{\beta\alpha}(M_{\alpha\alpha})^{-1} q_\alpha$	$M_{\beta\alpha}(M_{\alpha\alpha})^{-1}$	$M_{\beta\beta} - M_{\beta\alpha}(M_{\alpha\alpha})^{-1} M_{\alpha\beta}$

The DvPW Algorithm - Summary

	1	w_α	z_β
$q^\top z$	$-q_\alpha^\top (M_{\alpha\alpha})^{-1} q_\alpha$	$q_\alpha^\top (M_{\alpha\alpha})^{-1}$	$q_\beta^\top - q_\alpha^\top (M_{\alpha\alpha})^{-1} M_{\alpha\beta}$
z_α	$-(M_{\alpha\alpha})^{-1} q_\alpha$	$(M_{\alpha\alpha})^{-1}$	$-(M_{\alpha\alpha})^{-1} M_{\alpha\beta}$
w_β	$q_\beta - M_{\beta\alpha}(M_{\alpha\alpha})^{-1} q_\alpha$	$M_{\beta\alpha}(M_{\alpha\alpha})^{-1}$	$M_{\beta\beta} - M_{\beta\alpha}(M_{\alpha\alpha})^{-1} M_{\alpha\beta}$

- Maintain primal feasibility and strive for dual feasibility while maintaining complementary slackness.
- Choose $r \in \beta$ so that $q_r - M_{r\alpha}(M_{\alpha\alpha})^{-1} q_\alpha < 0$; if none, done.
- Increase z_r in each major cycle; this ends when w_r becomes nonbasic by being the blocking variable.
- Minor cycle occurs when one of the basic z_α-variables blocks the increase of z_r before w_r does.

CQP Results for the DvPW Algorithm

Theorem 1

Let $\kappa \triangleq \frac{4\lambda_{\max}(M)\|q_-\|_2^2}{\rho_{\min}(M)^2}$, *where* $q_- \triangleq \max\{0, -q\}$. *Then* $v_t - v_* \le \frac{\kappa n}{t-1}$.

* Note:

- $\lambda_{\max}(M)$ is the largest eigenvalue of M
- $\rho_{\min}(M) \triangleq \min_{\alpha : M_{\alpha\alpha} \succ 0} \lambda_{\min}(M_{\alpha\alpha})$, $\alpha \subseteq \{1, 2, ..., n\}$

* If $M \succ 0$, then $\rho_{\min}(M) = \lambda_{\min}(M)$.

To capture the finite termination property...

We need to introduce the two constants $\overline{\gamma}$ and $\underline{\delta}$ for the QP

CQP Results for the DvPW Algorithm

Theorem 2

If $M \succ 0$, then for any vector q, the DvPW algorithm with the least reduced cost rule computes the unique optimal solution of the QP in no more than

$$1 + 8\left(\frac{n\gamma_{\rm qp}}{\delta_{\rm qp}}\right)^2 cond(M)$$

iterations, where $cond(M) \triangleq \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)}$ is the condition number of M.

Key constants:

- $\delta_{\rm qp} \triangleq \min_{\text{feasible } \alpha} \min_{i\in\alpha} \left\{ z_i^\alpha \mid z_i^\alpha \triangleq \left[-(M_{\alpha\alpha})^{-1} q_\alpha\right]_i > 0 \right\}.$
- $\gamma_{\rm qp} \triangleq \max_{\text{feasible } \alpha} \max_{i\in\alpha} z_i^\alpha.$

CQP Results for the DvPW Algorithm

Theorem 2

If $M \succ 0$, then for any vector q, the DvPW algorithm with the least reduced cost rule computes the unique optimal solution of the QP in no more than

$$1 + 8\left(\frac{n\gamma_{\rm qp}}{\delta_{\rm qp}}\right)^2 cond(M)$$

iterations, where $cond(M) \triangleq \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)}$ is the condition number of M.

Key constants:

- $\delta_{\rm qp} \triangleq \min\limits_{\text{feasible } \alpha} \min\limits_{i\in\alpha} \left\{ z_i^\alpha \mid z_i^\alpha \triangleq \left[-(M_{\alpha\alpha})^{-1} q_\alpha \right]_i > 0 \right\}$.
- $\gamma_{\rm qp} \triangleq \max\limits_{\text{feasible } \alpha} \max\limits_{i\in\alpha} z_i^\alpha$.

Can we estimate them? Yes for certain problems.

Least Squares in Network Flow Problems

Consider a directed graph $G = (V, E)$ without self-loops, where $V = [n]$ is the set of vertices and $E \subseteq V \times V$ is the set of arcs. Assume $n \geq 2$. Let $A \in \mathbb{R}^{|V| \times |E|}$ denote the vertex-arc incidence matrix of G. For $v \in V$ and $e \in E$:

- $A_{ve} = 1$ if $e = (v, u)$ for some $u \in V$.
- $A_{ve} = -1$ if $e = (u, v)$ for some $u \in V$.
- $A_{ve} = 0$ otherwise.

We aim to solve the following problem:

$$\min_{x \geq 0} \ \frac{1}{2}\|Ax - b\|_2^2 + c^\top x, \tag{4}$$

where $b \in \mathbb{R}^n$ and $c \in \mathbb{R}_+^{|E|}$.
Note that *edge-Laplacian matrix* $M = A^\top A$ may not be positive definite unless G is a forest, Theorem (2) is not directly applicable.

Least Sqaures in Network Flow Problems Results

Proposition 1

When applying the DvPW algorithm to (4), one has

$$v_t - v_* \leq \frac{4n^6||b||_2^2}{\pi^2(t-1)}.$$

Proposition 2

Assume b and c are integer data. Then the DvPW algorithm with the least reduced cost rule computes the unique optimal solution of the QP (4) in no more than

$$\frac{n^4||b||_2^2}{2}$$

iterations.

Structured QP

Structured quadratic programs involve a matrix form:

$$M = K\Xi + FF^T.$$

Consider:

$$\min_{z \geq 0} \; z^\top \left(K\Xi + FF^\top \right) z + q^\top z, \tag{5}$$

where:

- $K > 0$ is an integer, $q \in \mathbb{R}^n$.
- Ξ is a positive definite matrix.
- $F \in \mathbb{R}^{n \times r}$ is a low-rank matrix.

Assumptions and context:

- All data are rational numbers; Ξ, F, q can be scaled to integers.
- Interested in: $\Xi_{\alpha\alpha}$ with small determinants and low-rank F.
- Example: factor models of portfolio risk analysis ($\Xi = I$, r is a small number of economic factors) [Atamtürk and Gómez, 2022, Bienstock, 1996, Konno and Suzuki, 1992].

Structured QP Results

Proposition 3

For integer data K, Ξ, F, and q with $D \triangleq \max_{\alpha \subseteq [n]} \det(\Xi_{\alpha\alpha})$, the DvPW algorithm computes the unique optimal solution of (5) in at most

$$1 + \frac{8n^2 C D^{2r+2} \|q\|_2^2}{K \lambda_{\min}(\Xi)^{2r+3}}$$

iterations, where $C = \left[K\lambda_{\min}(\Xi) + \lambda_{\max}(F)^2\right]^{2r} \left[K\lambda_{\max}(\Xi) + \lambda_{\max}(F)^2\right]$.

Iteration bound when $\Xi = I$:

$$1 + \frac{8n^2 \left[K + \lambda_{\max}(F)^2\right]^{2r+1} \|q\|_2^2}{K}.$$

Concluding Remarks

- Expand understanding of the efficiency of simplex-type methods
- Focus on classes of problems with favorable complexity bounds, in particular, strongly polynomial in problem magnitudes and derivable constants

Thank you for listening!

References I

Atamtürk, A. and Gómez, A. (2022).
Supermodularity and valid inequalities for quadratic optimization with indicators.
Mathematical Programming, pages 1–44.

Bienstock, D. (1996).
Computational study of a family of mixed-integer quadratic programming problems.
Mathematical programming, 74:121–140.

Cottle, R. W., Pang, J. S., and Stone, R. E. (2009).
The linear complementarity problem.
SIAM Publications.

Dantzig, G. (1963).
Linear programming and extensions.
Princeton University Press.

References II

Kitahara, T., Matsui, T., and Mizuno, S. (2012).
On the number of solutions generated by Dantzig's simplex method for LP with bounded variables.
Pacific Journal of Optimization, 8(3):447–455.

Konno, H. and Suzuki, K.-i. (1992).
A fast algorithm for solving large scale mean-variance models by compact factorization of covariance matrices.
Journal of the operations research society of Japan, 35(1):93–104.

Van de Panne, C. and Whinston, A. (1964).
The simplex and the dual method for quadratic programming.
Journal of the Operational Research Society, 15(4):355–388.

References III

Ye, Y. (2011).
The simplex and policy-iteration methods are strongly polynomial for the markov decision problem with a fixed discount rate.
Mathematics of Operations Research, 36(4):593–603.