

The Study of Finding an Approximately Optimal Rescaling in Context of Linear Programming

Krishna Dev Yadav¹, Dr. Om Prakash Dubey²

1. Research Scholar, P.G. Department of Mathematics, V. K. S. University, Ara

2. Assistant Professor, Department of Mathematics, J. J. College, Ara, V. K. S. University, Ara

email: krishnadevyadav.kd@gmail.com, omprakashdubeymaths@gmail.com

Abstract:

Our study area belongs to Linear Programming hence we have decided to study to find the approximately optimal rescaling in this paper. The paper first describes the standard primal and dual form of a constraint matrix. This paper also studies shortest path and interior-points method. By reviewing the earlier studies and research done in the field of linear programming, optimization, constraint matrix, optimal rescaling we are going to present our findings in this paper. Related theorems, propositions and corollaries seem to be part of our study and hence they are described in this paper. A result has been provided in this paper for a scaling-invariant Interior-Point Method that works out LP in time that is dimensionally a polynomial, answering an unlocked question by Monteiro-Tsuchiya in the assenting. This locks a long line of research by Vavasis-Ye and Monteiro-Tsuchiya;

Keywords: *Constraint matrix, approximation factor, optimal rescaling, circuit ratio diagram, Gaus-Jordan elimination.*

Introduction:

Linear Programming being a technique of optimizing the problems related to Operational Research. It is an analytical approach to formulate the problems related to Operational Research Development. Vavasis and Ye were the first to present the exact algorithm on the topic of Linear Programming in the real system model of computation. The running time of the algorithm was dependent on constraint matrix. By using the following standard primal and dual form for $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$

System 1: Linear Program	
Data: A triple (A, b, c) , in which $m \times n$ matrix is represented by A , $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.	
Primal	Dual
$\min_x (c, x)$ $\text{s. t. } Ax = b,$ $x \geq 0$	$\min_{y, s} (y, b)$ $\text{s. t. } A^T y + s = c,$ $s \geq 0$

One type of interior-point method was introduced by Vavasis and Ye that solves above represented linear programming related above represented primal and dual form. The iterations were $O(n^{3.5} \log(\bar{\chi}A + n))$, $\bar{\chi}A$ is the condition number and controls the size of the solutions to specified linear systems correlated to the kernel of A . We found that going further in this context of Vavasis and Ye (VY) algorithm it is better to discuss and describe the basics of path following interior-points method. The System represented above, if they are strictly feasible, the central path of it is curve $\{x(\mu), y(\mu), s(\mu): \mu > 0\}$ which is defined by

$$x(\mu)_i s(\mu)_i = \mu, \forall i \in [n]$$

$$Ax(\mu) = b, x(\mu) > 0,$$

$$A^T y(\mu) + s(\mu): \mu > 0,$$

The duality gap at time μ , that can be referred as normalized duality gap, is accurately $x(\mu)^T s(\mu) = n\mu$. The above relation converges to the presented optimal primal and dual solution x^*, y^*, s^* for earlier presented system. Also, the normalized duality gap can be defined by us as $\mu(x, y, s) = \sum_{i=1}^n x_i s_i / n$ while trying to achieve rapid multiplicative progress w.r.t. μ where certain (x, y, s) „close“ to the path for a method that “follow the path” generates iterates which continue in given neighbourhood around it. Given at parameter μ , starting point close to the path and an objective parameter μ' , in at most $O(\sqrt{n} \log(\mu/\mu'))$ iterations, standard path following methods can figure out a point at parameter lower than μ' , and so $\log(\mu/\mu')$ can be conveniently interpreted as the length of the equivalent segment of the central path.

Layered least squares steps and Crossover events

While conducting the study we found that Vavashi and Ye explain that the central path can be divided into at most $\binom{n}{2}$ curved short segments, possibly joined very straight and long segments. A new ordering relation termed as *crossover event* at the end of the each curved segment, $x_i(\mu) > x_j(\mu)$ is unconditionally discovered. This inequality guaranteed to hold at every point from the end of the segment onwards but it did not hold at the start of the segment. Along the central path, a combinatorial way to quantify progress is given by the $\binom{n}{2}$ relations. As per the analysis the cross-over event can only found to exist but not identified. This does not correlate the Tardos's algorithm variables are set to zero explicitly.

Description

Theorem 1: For a subspace $W \subset \mathbb{R}^n$, we have

$$\kappa_W^* = \min_{d>0} \kappa_W^d = \max\{\kappa_W(H)^{1/|H|}: H \text{ is a cycle in } G.\} \quad (1)$$

Proof: For the direction $\kappa_W(H)^{1/|H|} \leq \kappa_W^*$ we use

$$\kappa_W^d(H) = \kappa_W(H) \forall d \in \mathbb{R}_{++}^n \text{ for any cycle } H \text{ in } G \quad (2)$$

Let H being a cycle and $d > 0$ be a scaling. We have for every $j \in [n]$ $\kappa_{ij}^d \leq \kappa_W^d$, and therefore $\kappa_W(H) = \kappa_W^d(H) \leq (\kappa_W^d)^{|H|}$. As this holds for every $d > 0$, it follows that $\kappa_W(H) \leq (\kappa_W^*)^{|H|}$.

Here, we can consider the following optimization problem for the reverse direction

$$\begin{aligned} \min t \\ \kappa_{ij} d_j / d_i \leq t \quad \forall (i, j) \in E \\ d > 0 \end{aligned} \quad (3)$$

A feasible solution $(\lambda d, t)$ with the same objective value, is obtained for any feasible solution (d, t) and $\lambda > 0$. Without changing the objective value we can now strengthen the condition $d > 0$ to $d \geq 1$. Hence it is clear that the optimum value is realized by a feasible solution.

Any rescaling $d > 0$ presents a feasible solution with objective value κ^d . This means that represented by t^* of equation (3) is $t^* = \kappa^*$. $z_i = \log d_i, s = \log t$ when substituted, equation (3) can be written as a linear program:

$$\begin{aligned} \min s \\ \log \kappa_{ij} + z_j - z_i &\leq s \quad \forall (i, j) \in E \\ z &= \mathbb{R}^n \end{aligned} \quad (4)$$

The equation (4) is the dual of a minimum-mean cycle problem with respect to $\log(\kappa_{ij})$, the cost function. Hence, an optimal solution tallies to the maximizing $\kappa(H)^{1/|H|}$ or the cycle maximizing $\sum_{ij \in H} \log \kappa_{ij} / |H|$.

The theorem is thus proved but polynomial-time algorithm to compute κ_W^* is not obtained. The requirement is that the values κ_{ij}^W are typically not available. Their approximation up to a factor $2^{O(m)}$ is NP-hard, as the work of Tuncel is followed by it.

Further this theorem with a corollary, presented below, has been found to show that any arbitrary circuit containing i and j generates a $(\kappa^*)^2$ approximation to κ_{ij} .

Corollary 1: Having a linear subspace $W \subseteq \mathbb{R}^n$ and $i, j \in [n], i \neq j$, and with $i, j \in C$ a circuit $C \in \mathcal{C}$, Assuming $\mathcal{G} \in W$ be the corresponding vector with $\text{supp}(\mathcal{G}) = C$, then

$$\frac{\kappa_{ij}^W}{(\kappa_W^*)^2} \leq \frac{|\mathcal{G}_j|}{|\mathcal{G}_i|} \leq \kappa_{ij}^W$$

Proof: The definition is found to correlate with second inequality. The same circuit C generates $|\mathcal{G}_i/\mathcal{G}_j| \leq \kappa_{ij}^W(C) \leq \kappa_{ji}^W$ for the first inequality. So, it can be concluded that $|\mathcal{G}_i/\mathcal{G}_j| \geq 1/\kappa_{ji}^W$. From above theorem it is clear that $\kappa_{ij}^W \kappa_{ji}^W \leq (\kappa_W^*)^2$, giving $1/\kappa_{ji}^W \geq \kappa_{ij}^W / (\kappa_W^*)^2$, completing the proof.

With the use of above theorem and corollary we have tried to identify a circuit for any pair of variables by using techniques from linear algebra and matroid theory that are contained in the same circuit.

Theorem 2: There exists an $O(n^2 m^2)$ time algorithm *FIND – CIRCUITS* (A) for a given $A \in \mathbb{R}^{m \times n}$ that to a direct sum of non-separable linear matroids obtains a decomposition of $\mathcal{M}(A)$ and returns a family $\mathcal{C}^\wedge$ of circuits such that there exists a circuit $C^\wedge$ containing both i and j when i and j are in the same non-separable component. Further, the algorithm returns a value $\kappa_{ij}^\wedge$ for each $i \neq j$ in the same component as the maximum of $|\mathcal{G}_j/\mathcal{G}_i|$ such that for some $C \in \mathcal{C}^\wedge$ containing i and $\mathcal{G} \in W, \text{supp}(\mathcal{G}) = C$. For these values, $\kappa_{ij}^\wedge \leq \kappa_{ij} \leq (\kappa^*)^2 \kappa_{ij}^\wedge$.

Proof: Once we have computed $\kappa_{ij}^{\wedge}$ as in the statement having the set of circuits $C^{\wedge}$, the inequalities $\kappa_{ij}^{\wedge} \leq \kappa_{ij} \leq (\kappa^*)^2 \kappa_{ij}^{\wedge}$ follow easily. κ_{ij} defines the first inequality, and the second inequality is defined by a corollary 1.

For the computation of $C^{\wedge}$, we first obtain a basis of $\ker(\mathbf{A})$ as $B \subseteq [n]$ via Gauss-Jordan elimination in time $O(nm^2)$. The assumption that $\mathbf{A}$ has full row-rank made us during the study to assume that $B = [m]$ is the set of first m indices. $H \in \mathbb{R}^{m \times (n-m)}$ which corresponds to the non-basis-elements helped us to find that the elimination transforms to the basis form $\mathbf{A} = [\mathbf{I}_m \quad \mathbf{H}]$. The fundamental circuit $C(B, i)$ in this form is the support of the i -th column of $\mathbf{A}$ in addition to i for every $m + 1 \leq i \leq n$. We let C^B designate the set of all such elementary circuits.

An undirected graph $G = (B, E)$ is constructed by us in $O(nm^2)$ time as for each $i \in [n] \setminus B$, a clique between the nodes in $C(B, i) \setminus \{i\}$ is added.

The connected component of constructed graph correspond to the linked components of C^B restricted to B . We further can, having some proposition in this context, obtain the decomposition by recognizing the connected components of graph G . By assuming that the entire hypergraph is connected; in time $O(m^2)$ the connectivity can be checked so that rest part of the theorem can be proved by us.

Initialize $C^{\wedge}$ as C^B

Check all pairs $i, j \in [n], i \neq j$

A circuit $C^{\wedge}$ as follows is required to added when no circuit $C \in C^{\wedge}$ exists with $i, j \in C$.

Let $i, j \in [n] \setminus B$. Between the sets $C(B, i) \setminus \{i\}$ and $C(B, j) \setminus \{j\}$ in time $O(m^2)$ the shortest path in Graph G is obtained. The obtained shortest path thus can be represented by the series of points $V = \{v_1, v_2, \dots, v_{t+1}\} \subseteq [n] \setminus B, v_i = i, v_{t+1} = j$ and $U = \{u_1, u_2, \dots, u_t\} \subseteq B$. Now we can have $S = (B \setminus U) \cup V$ that contains a unique circuit C that further contains all v_t 's, together with i and j .

To show how this circuit can be recognized in time $O(m)$ in conjunction with the vector g^C , let $\mathbf{A}_S$ be the submatrix in line with the columns of S . Since $g = g^C$ is distinctive up to scaling, having $g_{v_1} = 1$. Suggested that for each $t \in [\ell]$, the row of $\mathbf{A}_S$ along the line of u_t contains only two nonzero entries: $A_{u_t v_t}$ and $A_{u_t v_{t+1}}$. Hence $g_{v_1} = 1$ can be propagated to initialising unique values to $g_{v_2}, g_{v_3}, \dots, g_{v_{t+1}}$. After setting these values, there is a unique extension of g to indices $t \in B \cap S$ in the basis. Here g has been identifies as the distinctive element of $\ker(\mathbf{A}_S)$ up to scaling. $\text{supp}(g)$ is the obtained circuit C . The entire procedure can be executed in time $O(m)$.

Theorem 3: For any matrix $A \in \mathbb{R}^{m \times n}$ we have $\sqrt{1 + \kappa_A^2} \leq \bar{\chi}_A \leq n\kappa_A$.

Proof: For the corresponding solution $g = g^C$ let for the first inequality $C \in C_W$ be the circuit and $i \neq j \in C$ such that $|g_j / g_i| = \kappa_W$. Let us use the characterization of $\bar{\chi}_W$ in the following proposition

Proposition 1: For a linear subspace $W \subseteq \mathbb{R}^n$,

$$\bar{\chi}_W = \max\{||L^W|| : I \subseteq [n], I \neq \emptyset\}.$$

The above proposition has been used to represent the proof of Theorem 1 and hence is not a matter of discussion in this paper, so it has been left. Now, let $I = ([n] \setminus C) \cup \{i\}$ and the vector $p_i = g_i$ and $p_k = 0$ for $k \neq i$, $p = g_i e^i$, and so the unique vector $z \in W$ such that $z_i = p$ is $z = g$. Hence,

$$\bar{\chi}_W \geq \min_{z \in W, z_i = p} \frac{\|z\|}{\|p\|} = \frac{\|g\|}{|g_i|} \geq \frac{\sqrt{|g_i|^2 + |g_j|^2}}{|g_i|} = \sqrt{1 + \kappa_W^2}$$

For further description of above result, another proposition is required here as:

Proposition 2: For a linear subspace $W \subseteq \mathbb{R}^n$,

$$\kappa_W = \max\left\{\frac{\|L_I^W(p)\|}{\|p\|_I} : I \subseteq [n], I \neq \emptyset, p \in \pi_I(W) \setminus \{0\}\right\}$$

and the Lemma as:

Given a linear subspace $W \subseteq \mathbb{R}^n$, a threshold $\theta \in \mathbb{R}_{++}$ and an index set $I \subseteq [n]$ there exists a subroutine *VERIFY – LIFT*(W, I, θ) that on verifying $\ell^W(I) \leq \theta$, either returns an answer „pass“ or returns „fail“ as answer, and a pair $i \in I, j \in [n] \setminus I$ such that $\theta/n \leq \kappa_{ij}^W$. The running time of the subroutine can be bounded as $O(n(n - m)^2)$.

The inequality represented above is instantaneous from Proposition 1 and 2 and the inequalities ℓ_1 , ℓ_2 and ℓ_∞ norms. The proof of somewhat weaker $\bar{\chi}_W \leq \sqrt{1 + (n\kappa_W)^2}$ pursues from used lemma here. Some other lemma may appear to be proved but this is not the too much required to describe them in this paper.

Theorem 4: For any matrix $A \in \mathbb{R}^{m \times n}$ there is an $O(n^2 m^2 + n^3)$ time algorithm that determines an estimate ξ of $\bar{\chi}_W$ such that

$$\xi \leq \bar{\chi}_W \leq n(\bar{\chi}_W)^2 \xi$$

and $aD \in \mathcal{D}$ such that

$$\bar{\chi}_W \leq \bar{\chi}_{DW} \leq n(\bar{\chi}_W)^3$$

Proof: The Theorem 2 is executed to obtain the values κ_{ij}^A such that $\kappa_{ij}^A \leq \kappa_{ij} \leq (\kappa_W^*)^2 \kappa_{ij}^A$. Let $G = ([n], E)$ be the circuit ratio diagram, $(i, j) \in E$ if $\kappa_{ij} > 0$.

$\xi = \max_{(i,j) \in E} \hat{\kappa}_{ij}$ is set to show the first statement of approximating $\bar{\chi}$, and we have

$$\xi \leq \kappa_W \leq \bar{\chi}_W \leq n\kappa_W \leq n(\kappa_W^*)^2 \xi \leq n(\bar{\chi}_W^*) \xi$$

follows theorem 1.

Now we have to consider the following optimization problem for the second statement on evaluating an optimal rescaling for $\bar{\chi}_W$, which is an approximate optimization problem of (3) from theorem 1.

$$\begin{aligned} \min t \\ \kappa_{ij} \hat{d}_j / d_i \leq t \quad \forall (i, j) \in E \\ d > 0 \end{aligned} \tag{5}$$

Assume d^* be an optimal solution to (5) having value t^* . We will now prove that $\kappa^d \leq (\kappa_W^*)^3$.

First for any $(i, j) \in E$ observe that $\kappa^d = \frac{\kappa_{ij} \hat{d}_j}{d_i^*} \leq \frac{(\kappa_W^*)^2 \hat{\kappa}_{ij} \hat{d}_j}{d_i^*} \leq (\kappa_W^*) t^*$. Now, Assume, $d^* > 0$ be such that $\kappa^{d^*} = \kappa_W^*$. The vector d^* appears to be a feasible solution to (5), and so $t^* \leq \max_{i \neq j} \kappa_{ij} d^* / d_i^* \leq \max_{i \neq j} \kappa_{ij} d^* / d_i^* = \kappa^{d^*}$. Hence we obtain that d^* gives a rescaling with

$$\bar{\chi}_{W^d} \leq n\kappa^d \leq n(\kappa_W^*)^3 \leq n(\bar{\chi}_W)^3,$$

where theorem 3 is again used by us.

The optimal value of t^* of (5) is now obtained by us by solving the corresponding maximum-mean cycle problem (Theorem 1). A multiplicative form of the standard dynamic programming algorithm of the classical minimum-mean cycle problem, can be easily developed that allows evaluating the optimum to (5) in the same time $O(n^3)$ directly.

Now to find the labels $d_i > 0, i \in [n]$ such that $\kappa_{ij} \hat{d}_j / d_i \leq t^*$ for all $(i, j) \in E$ let us define the following weighted directed graph. The weight $w_{ij} = \log t^* - \log \kappa_{ij}$ with every $(i, j) \in E$ is associated by us and an extra source vertex r is added with edges (r, i) of weight $w_{ri} = 0$ for all $i \in [n]$.

If t^* is chosen, this graph does not enclose any negative weighted directed cycles. The shortest path from r to all nodes in $O(n^3)$ can also be computed by us using Bellman-Ford algorithm; let the shortest path label for i be σ_i . We then set $d_i = \exp(\sigma_i)$.

$O(n^2m^2 + n^3)$ bounds the running time of whole algorithm. The running time for complexity of *FINDING - CIRCUIITS*(A) is dominated by the $O(n^2m^2)$ and the complexity of solving the minimum mean cycle problem and shortest path computation is $O(n^3)$.

Literature Review:

Dikin, Stewart and Todd were first to study condition number. A large number of researchers have studied and presented their literature on the properties and applications of $\bar{\chi}_A$ and its relation to other condition number. It plays an important role in layered-least-squares interior point method. For preparing the content of this paper some other related papers have also been studied.

Findings:

Our finding in this paper is that we given an $O(m^2n^2 + n^3)$ time algorithm. To compute a diagonal rescaling matrix $\mathbf{D}$ which attains $\bar{\chi}_{AD} \leq n(\bar{\chi}_A^*)^2$ for any $m \times n$ matrix the obtained algorithm works. It also allows us to approximate $\bar{\chi}_A$ to within a factor $n(\bar{\chi}_A)^2$, which increases the runtime only by a constant factor. Allowing the approximation factor to depend on $\mathbf{A}$ itself, $(\bar{\chi}_A^*)$, this algorithm bypasses Tuncel's hardness results. Vavasis-Ye algorithm is applied directly on the preprocessed $\mathbf{A}$ matrix and hence the Monteiro-Tsuchiya's question has been answered. Thus we can state that we may solve any LP with constraint matrix $\mathbf{A}$ using $O(n^{3.5} \log(\bar{\chi}_A^* + n))$ iterations.

Conclusion:

Moneteiro and Tsuchiya, noting that under the rescalings of the columns of $\mathbf{A}$ and c , noted that the central path is invariant and presented a question before the researchers whether there exists and LP algorithm depending as an alternative on the measure $\bar{\chi}_A^*$, described as the minimum $\bar{\chi}_{AD}$ value feasible by a column rescaling AD of $\mathbf{A}$. Near-optimal rescaling of the constraint matrix is provided in this paper to answer the question.

References:

1. I. Dikin, "Iterative solution of problems of linear and quadratic programming", *Doklady Akademii Nauk*, 174:4, 1967, pp. 747-748,
2. G. Stewart, "On scaled projections and pseudoinverses", *Linear Algebra and its Applications*, 112, 1989, pp. 189-193
3. M. J. Todd, "A Dantzig-Wolfe-Like Variant of Karmarkar's Interior-Point Linear Programming Algorithm", *Operations Research*, 38:6, 1990, pp 106-1018, url: <https://doi-org/10.1287/opre-38.6.1006>
4. S. A. Vavasis and Y. Ye., "A primal-dual interior point method whose running time depends only on the constraint matrix, *Mathematical Programming*, 74:1, 1996, pp. 79-120.
5. L. Tuncel, "Approximating the complexity measure of Vavasis-Ye algorithm in NP-hard", *Mathematical Programming*, 86:1, 1999, pp. 219-223
6. R. D. C. Monteiro and T. Tsuchiya, "A modified layered-step interior-point algorithm for linear programming", *SIAM Journal on Optimization*, 13:4, 2003, pp. 1054-1079.
7. R. D. C. Monteiro and T. Tsuchiya, "A New Iteration-Complexity Bound for the MTY Predictor-Corrector Algorithm", *SIAM Journal on Optimization*, 15:2, 2005, pp. 319-347
8. R. D. C. Monteiro and T. Tsuchiya, "A strong bound on the integral of the central path curvature and its relationships with the iteration-complexity of primal-dual path-

- following LP algorithms”, *Mathematical Programming*, 115:1, 2008, pp. 105-149.
9. S. Bordwardt, T. Stephen and T. Yusun, “On the circuit diameter conjecture”, *Discrete & Computation Geometry*, 60:3, 2018, pp. 558-587
 10. J. van den Brand, “A Deterministic Linear Program Solver in Current Matrix Multiplication Time”, In: *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2020
 11. J. van den Brand, Y. P. Liu, Y. –T. Saranurak, A. Sidford, Z. Song and D. Wang., “Minimum Cost Flows, MDPs, and L1-Regression in Nearly Linear Time for Dense Instances”, In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, 2021, pp. 859 - 869