Applications of interior point methods for very large problems arising in imaging and optimal transport

Filippo Zanetti Jacek Gondzio
Samuli Siltanen Salla Latva-Aijo Matti Lassas

University of Edinburgh, School of Mathematics

Modern Techniques of Very Large Scale Optimization
20 May 2022

1. Interior point method introduction

2. Tomographic imaging applications

3. Optimal transport applications

Interior point method

Pair of primal-dual convex QPs

$$\min_{x}\ c^T x + \frac{1}{2} x^T Q x, \qquad \text{s.t. } Ax = b, \qquad x \geq 0,$$

$$\max_{y,\ s}\ b^T y - \frac{1}{2} x^T Q x, \qquad \text{s.t. } A^T y + s - Qx = c, \qquad s \geq 0,$$

$x, s, c \in \mathbb{R}^n$, $y, b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$, $Q \in \mathbb{R}^{n\times n}$ positive semidefinite.

An IPM enforces the non-negativity constraint using a **logarithmic barrier**.

At each iteration, a direction is computed that allows the approximation to get closer to **optimality** and **feasibility**.

Newton system

The search direction is found using the **Newton** method

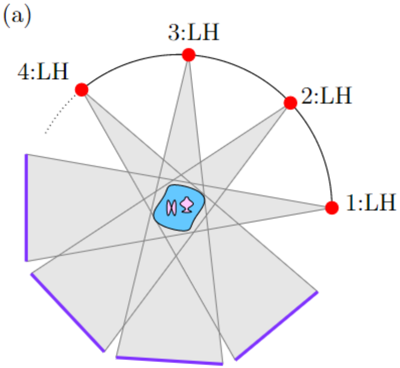
$$\begin{bmatrix} A & 0 & 0 \\ -Q & A^T & I \\ S & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} r_P \\ r_D \\ r_\mu \end{bmatrix} = \begin{bmatrix} b - Ax \\ c + Qx - A^T y - s \\ \sigma\mu e - XSe \end{bmatrix}.$$

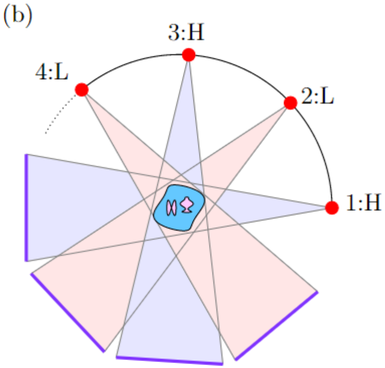
The system can be reduced to the **Normal equations**

$$A(Q+\Theta^{-1})^{-1}A^T\Delta y \;=\; r_P + A(Q+\Theta^{-1})^{-1}(r_D - X^{-1}r_\mu)$$

where $\Theta = XS^{-1}$, and solved using a factorization or the PCG.

Tomographic imaging





Tomographic imaging

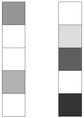
Problem formulation

$$\min_{x\geq 0} \quad \underbrace{\|m - \mathcal{G}x\|^2}_{\substack{\text{Cost}\\ \text{function}}} \quad + \quad \underbrace{\alpha\|x\|^2}_{\substack{\text{Tikhonov}\\ \text{regularizer}}} \quad + \quad \underbrace{2\beta\, x_1^T x_2}_{\substack{\text{Inner product}\\ \text{regularizer}}}$$

Tomographic imaging

Problem formulation

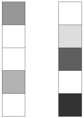
$$\min_{x \geq 0} \underbrace{\|m - \mathcal{G}x\|^2}_{\substack{\text{Cost}\\ \text{function}}} + \underbrace{\alpha\|x\|^2}_{\substack{\text{Tikhonov}\\ \text{regularizer}}} + \underbrace{2\beta\, x_1^T x_2}_{\substack{\text{Inner product}\\ \text{regularizer}}}$$



Tomographic imaging

Problem formulation

$$\min_{x\geq 0} \quad \underbrace{\|m-\mathcal{G}x\|^2}_{\text{Cost function}} \quad + \quad \underbrace{\alpha\|x\|^2}_{\text{Tikhonov regularizer}} \quad + \quad \underbrace{2\beta\, x_1^T x_2}_{\text{Inner product regularizer}}$$



$$\Downarrow$$

QP formulation

$$\min_{x\geq 0} \ \frac{1}{2}x^T\mathcal{Q}x - m^T\mathcal{G}x, \tag{1}$$

where

$$\mathcal{Q} = \begin{bmatrix} c_{11}^2 + c_{21}^2 & c_{11}c_{12} + c_{21}c_{22} \\ c_{11}c_{12} + c_{21}c_{22} & c_{12}^2 + c_{22}^2 \end{bmatrix} \otimes R^TR + \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \otimes I,$$

$c_{11}, c_{12}, c_{21}, c_{22} > 0$, $\alpha > \beta > 0$.

R is accessed through matrix-vector products performed using the **Radon transform**.

IPM formulation

Lemma

If $\alpha \geq \beta$, problem (1) is convex.

Applying an IPM to this problem, we need to solve:

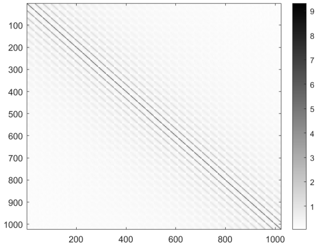
$$(\mathcal{Q} + X^{-1}S)\Delta\mathbf{x} = \mathbf{f}$$

This can be done using the Preconditioned Conjugate Gradient (PCG).

Gondzio, Lassas, Latva-Aijo, Siltanen, Z. *Material-separating regularizer for multi-energy X-ray tomography*. Inverse Problems, 38 (2022)

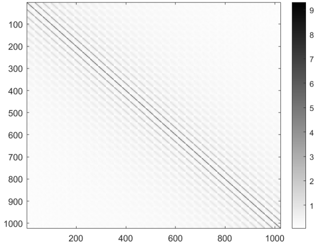
Preconditioner

$$\mathcal{Q} + X^{-1}S \quad = \quad \begin{bmatrix} c_{11}^2 + c_{21}^2 & c_{11}c_{12} + c_{21}c_{22} \\ c_{11}c_{12} + c_{21}c_{22} & c_{12}^2 + c_{22}^2 \end{bmatrix} \otimes \underbrace{R^T R} + \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \otimes I + X^{-1}S$$



Preconditioner

$$\mathcal{Q} + X^{-1}S = \begin{bmatrix} c_{11}^2 + c_{21}^2 & c_{11}c_{12} + c_{21}c_{22} \\ c_{11}c_{12} + c_{21}c_{22} & c_{12}^2 + c_{22}^2 \end{bmatrix} \otimes \underbrace{R^T R} + \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \otimes I + X^{-1}S$$



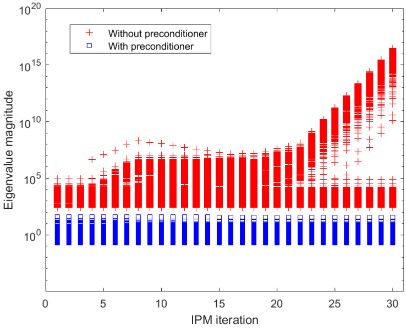
$$\mathcal{P} = \begin{bmatrix} c_{11}^2 + c_{21}^2 & c_{11}c_{12} + c_{21}c_{22} \\ c_{11}c_{12} + c_{21}c_{22} & c_{12}^2 + c_{22}^2 \end{bmatrix} \otimes \underbrace{\rho I} + \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix} \otimes I + X^{-1}S$$

Preconditioned matrix

Lemma

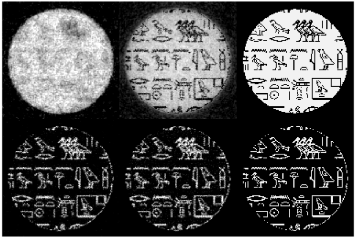
The eigenvalues of the preconditioned matrix $\mathcal{P}^{-1}(\mathcal{Q}+X^{-1}S)$ are independent of the IPM iteration and satisfy

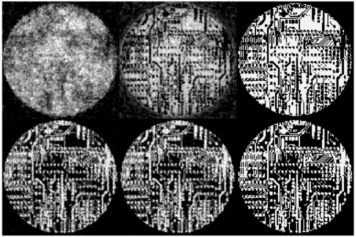
$$\lambda \in \left[\frac{\alpha-\beta}{\rho\Lambda_F+\alpha+\beta}, \frac{\sigma^2_{max}(R)\Lambda_F+\alpha+\beta}{\rho\lambda_F+\alpha-\beta}\right]$$



Results

JTV **IP** **Ground truth**





Discrete optimal transport

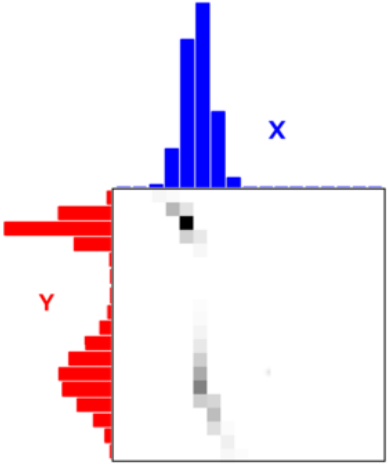


Image from en.wikipedia.org/wiki/Transportation_theory_(mathematics)

Problem formulation

Optimal transport formulation

$$\min_{\substack{\mathcal{P}\in\mathbb{R}_+^{m\times m}\\ \mathcal{P}\mathbf{e}_m=\mathbf{a},\ \mathcal{P}^T\mathbf{e}_m=\mathbf{b}}} \sum_{i,j} \mathcal{C}_{ij}\mathcal{P}_{ij}$$

Problem formulation

Optimal transport formulation

$$\min_{\substack{\mathcal{P}\in\mathbb{R}_+^{m\times m}\\ \mathcal{P}\mathbf{e}_m=\mathbf{a},\ \mathcal{P}^T\mathbf{e}_m=\mathbf{b}}} \quad \sum_{i,j} \mathcal{C}_{ij}\mathcal{P}_{ij}$$

$$\Downarrow$$

Kantorovich Linear Program

$$\begin{aligned} \min_{\mathbf{p}\in\mathbb{R}^{m^2}} \quad & \mathbf{c}^T\mathbf{p} \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{e}_m^T \otimes I_m \\ I_m \otimes \mathbf{e}_m^T \end{bmatrix} \mathbf{p} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \mathbf{f} \\ & \mathbf{p} \geq 0 \end{aligned}$$

IPM formulation - 1

Highly structured constraint matrix

$$\begin{bmatrix}
1 & & & & & 1 & & & & & 1 & & & & & \\
 & 1 & & & & & 1 & & & & & 1 & & & & \dots \\
 & & 1 & & & & & 1 & & & & & 1 & & & \\
 & & & 1 & & & & & 1 & & & & & 1 & & \\
 & & & & \ddots & & & & & \ddots & & & & & \ddots & \\
1 & 1 & 1 & 1 & \dots & & & & & & & & & & & \\
 & & & & & 1 & 1 & 1 & 1 & \dots & & & & & & \dots \\
 & & & & & & & & & & 1 & 1 & 1 & 1 & \dots & \\
 & & & & & & & & \vdots & & & & & & &
\end{bmatrix}.$$

- $2m$ rows
- m^2 columns

IPM formulation - 2

Solved with a sparse mixed IPM-Column Generation approach:

- Keeps the iterates as **sparse** as possible
- Exploits **second order** method convergence

Linear system solved using the Schur complement of the normal equations:

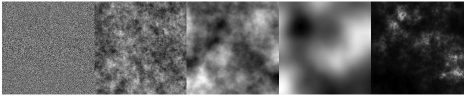
- PCG with **incompete Cholesky** factorization in the early stage
- Full LDL^T **factorization** in the late IPM iterations

Highly optimized **matrix-free sparse implementation**

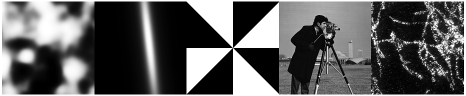
A full IPM would involve a **dense** Schur complement and very **large** and dense vectors

Test problems

DOTmark collection



Class 1 Class 2 Class 3 Class 4 Class 5



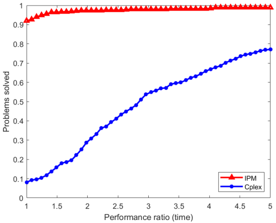
Class 6 Class 7 Class 8 Class 9 Class 10

res	Constraints	Variables ($\times 10^6$)
32	2,048	1.0
64	8,192	16.8
128	32,768	268.4
256	131,072	4,295.0

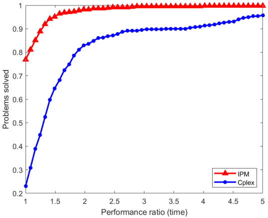
Cost functions: $1-$norm, $2-$norm, $\infty-$norm.

Results

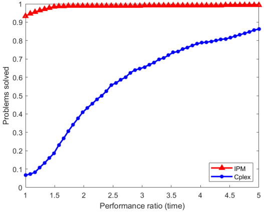
Comparison of **SparseIPM** and IBM ILOG Cplex **network simplex** solver



1−dist, 64 × 64



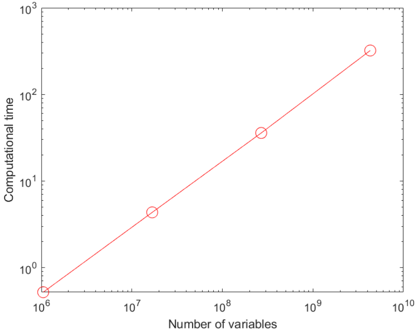
2−dist, 64 × 64



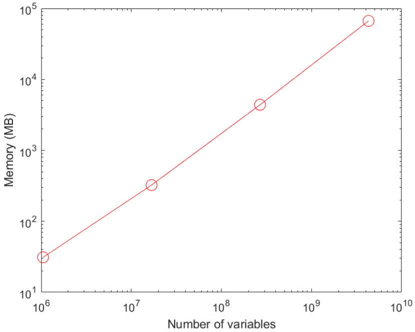
∞−dist, 64 × 64

Large instances results

Scalable results up to **4 Billion** variables (DOTmark class 1 problems)



Time



Memory

Thanks for the attention