KTH ROYAL INSTITUTE
OF TECHNOLOGY

Interior-point method for nuclear norm approximation with application to system identification

Authors : Zhang Liu and Lieven Vangenberghe

Presented by Ezgi Korkmaz

SIAM Journal on Matrix Analysis and Applications (SIMAX) publishes research papers on,

- Matrix and tensor theory,
- Analysis, applications
- Computation that are of interest to the applied and numerical linear algebra

Applications include such areas as signal processing, systems and control theory, statistics, Markov chains, and mathematical biology.

The Problem

- The problem of minimizing the nuclear norm of an affine matrix valued function
- Can be formulated as semidefinite programming
- Authors show that semidefinite programming formulation can be exploited to develop more efficient interior-point methods.
- Cost per iteration reduced to a quartic function of the problem dimensions
- The convex nuclear norm heuristic has applications on
  - Control and system theory in model reduction and system identification
  - Machine learning and computer vision
  - Recommender systems

Formulation as an optimization problem

Implementation of interior-point algorithms for the nuclear norm approximation

$$\text{minimize}||A(x) - B||_* \tag{1}$$

$B \in R^{p \times q}$ is a given matrix and
$A(x) = x_1 A_1 + x_2 A_2 + ... + x_n A_n$ is a linear mapping from R^n to $R^{p \times q}$.

$$||X||_* = \sum \sigma_i(X) = trace(\sqrt{(X * X)}) \tag{2}$$

Also applicable to problems with an added convex quadratic term in the objective

$$\text{minimize}||A(x) - B||_* + (x - x_0)^T Q(x - x_0) \tag{3}$$

Rank minimization

Nuclear norm approximation is interest of as a convex heuristic for the rank minimization problem,

$$\text{minimize}\ \operatorname{rank}(A(x) - B) \tag{4}$$

which is NP hard in general.
Application areas,

- Model reduction, system identification, minimum order control synthesis in control
- Minimum rank completion problem in machine learning
- Nonnegative factorization in data mining

The nuclear norm heuristic has been observed to produce very low-rank solutions in practice

Formulation as an optimization problem

$$\text{minimize} ||A(x) - B||_* \tag{5}$$

$$B \in R^{p \times q} \tag{6}$$

$$A(x) = x_1 A_1 + x_2 A_2 + ... + x_n A_n \tag{7}$$

is convex,

$$||A(\alpha x + (1-\alpha) - B||_* = ||\alpha(A(x)-B) + (1-\alpha)(A(y)-B)|| \leq$$
$$||\alpha(A(x) - B)||_* + ||(1-\alpha)(A(y) - B)||_* =$$
$$\alpha||A(x) - B||_* + (1-\alpha)||A(y) - B||_*$$

Formulation as an optimization problem

The nuclear norm approximation can be cast as SDP,

$$\begin{array}{ll} \text{minimize} & (\operatorname{tr} U + \operatorname{tr} V)/2 \\ \text{subject to} & \begin{bmatrix} U & (A(x) - B)^T \\ A(x) - B & V \end{bmatrix} \geq 0 \end{array}$$

Interior methods offer fast convergence; however, the volume of computation per iteration is high.
The main contribution of the paper solving the problem with dimensions p,q on the order of several hundreds while SDP having problem when p and q reaches 100.

Custom interior-point method

The nuclear norm approximation can be cast as SDP,

$$\begin{array}{ll} \underset{x}{\text{minimize}} & (\operatorname{tr} U + \operatorname{tr} V)/2 \\ \text{subject to} & \begin{bmatrix} U & (A(x) - B)^T \\ A(x) - B & V \end{bmatrix} \geq 0 \end{array}$$

The solution of a set of linear equations,

$$A_{adj}(\Delta Z) = r, \begin{bmatrix} \Delta U & A(\Delta x)^T \\ A(\Delta x) & \Delta V \end{bmatrix} + T \begin{bmatrix} 0 & (\Delta Z)^T \\ \Delta Z & 0 \end{bmatrix} T = R \tag{8}$$

T is positive semidefinite, R is the residuals. To compute search directions $\Delta x, \Delta U, \Delta V$ and ΔZ

Custom interior-point method

$$\begin{array}{ll} \underset{x}{\text{minimize}} & (\operatorname{tr} U + \operatorname{tr} V)/2 \\ \text{subject to} & \begin{bmatrix} U & (A(x) - B)^T \\ A(x) - B & V \end{bmatrix} \geq 0 \end{array}$$

The solution of a set of linear equations,

$$A_{adj}(\Delta Z) = r, \begin{bmatrix} \Delta U & A(\Delta x)^T \\ A(\Delta x) & \Delta V \end{bmatrix} + T \begin{bmatrix} 0 & (\Delta Z)^T \\ \Delta Z & 0 \end{bmatrix} T = R \tag{9}$$

can be solved by first finding the Δx

$$A_{adj}(\mathcal{S}^{-1}(A(\Delta X))) = A_{adj}(S^{-1}(R_{21})) - r \tag{10}$$

Custom interior-point method

Solving Δx looks complicated.

$$\mathcal{A}_{adj}(\mathcal{S}^{-1}(\mathcal{A}(\Delta X))) = \mathcal{A}_{adj}(\mathcal{S}^{-1}(R_{21})) - r \tag{11}$$

However,

$$\mathcal{S}(X) = \mathcal{L}(\mathcal{L}_{adj}(X)) \tag{12}$$

where the inverses of $\mathcal{L}$ are defined as,

$$\mathcal{L}^{-1}(X)_{ij} = \begin{cases} X_{ij} - \sigma_i \sigma_j X_{ji} / \sqrt{(1 - \sigma_i^2 \sigma_j^2} & i < j \\ X_{ii} / \sqrt{1 + \sigma_i^2} & i = j \\ X_{ij} & i > j \end{cases}$$

Complexity

The linear algebra complexity per iteration of the custom interior-point algorithm can be estimated as follows,

- The cost of computing coefficients of the scaled mapping, $A(x) = x_1 A_1 + x_2 A_2 + \ldots + x_n A_n$ is $\mathcal{O}(np^2 q)$
- Solving by Cholesky factorization costs $\mathcal{O}(n^2 pq)$
- More generally, if $p = \mathcal{O}(n)$ and $q = \mathcal{O}(n)$
    - The cost per iteration increases as $\mathcal{O}(n^4)$ as opposed to $\mathcal{O}(n^6)$

Numerical Experiments

Complexity reduced from $\mathcal{O}(p^2q^2n)$ to $\mathcal{O}(pqn^2)$

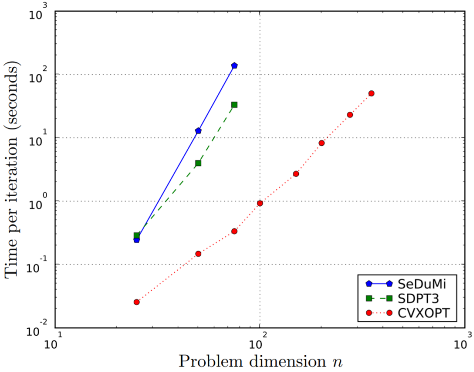

Figure: Numerical experiment results

An Open Loop Control Problem

$$\begin{aligned} \text{minimize} \quad & ||\mathcal{H}(x)||_* \\ \text{subject to} \quad & b_l(t) \leq (h * x * u)(t) \leq b_u(t), t = 0, ...N \end{aligned}$$

b_l and b_u denotes the upper and lower bound of the step response.

$$\mathcal{H}(x) = \begin{bmatrix} x(1) & x(2) & x(3) & ... & x(N) \\ x(2) & x(3) & x(4) & ... & x(N+1) \\ . & . & . & ... & . \\ . & . & . & ... & . \\ x(N) & x(N+1) & x(N+2) & ... & x(2N-1) \end{bmatrix} \tag{13}$$

h(t), the plant impulse response, x(t) the controller impulse response and u(t) is the step input

An Open Loop Control Problem

The results for 6th order plant with transfer function

$$P(z) = \frac{0.015z^5 - 0.00292z^4 - 0.0284z^3 + 0.0177z^2 + 0.00816z - 0.00828}{z^6 - 4.03z^5 + 7.4z^4 - 8.06z^3 + 5.57z^2 - 2.31z + 0.434} \quad (14)$$

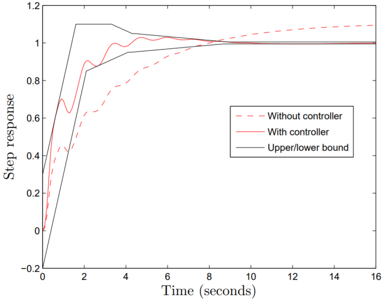

Figure: Open loop controller design with time domain constraints

- Efficient and useful algorithm
- Implementation is tedious
- Unclear comments on numeraical results

Thank you!