KATHOLIEKE UNIVERSITEIT LEUVEN

OPTEC NALAG

Optimization of Large-scale Systems with Model Reduction

Presenter: Yao Yue

Promotor: Karl Meerbergen

September 17, 2009

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

Acoustic Cavity Design

A cubic acoustic cavity:

- One side: Admittance boundary condition.
- Other sides: Dirichlet boundary condition.
- Design Parameter: Admittance ratio.
- Design Objective: Minimize the highest peak of FRF at a given point.

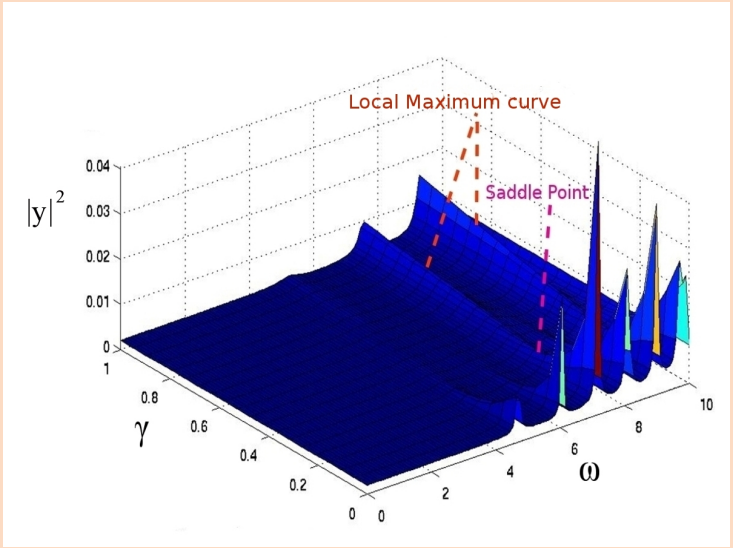
Mathematical Formulation

$$\min_{\gamma} \max_{\omega_L \le \omega \le \omega_H} |y(\omega,\gamma)|^2$$

$$\text{s.t.} \quad \begin{cases} y(\omega,\gamma) = \ell^* x(\omega,\gamma) \\ (K + i\omega\gamma C - \omega^2 M)x(\omega,\gamma) = f \end{cases}$$

Minimax Optimization

3D Figure



All local maximum curves are convex!

Optimizer

Local optimizer

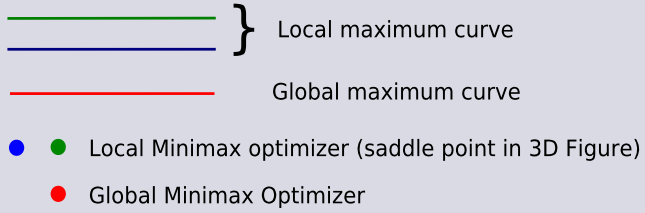
$\Longleftrightarrow$ minimum of local maximum curve

$\Longleftrightarrow$ Saddle Point.

What about the global optimizer?

Global Optimizer

Project the 3D Figure to $|y|^2 - \gamma$ plane.



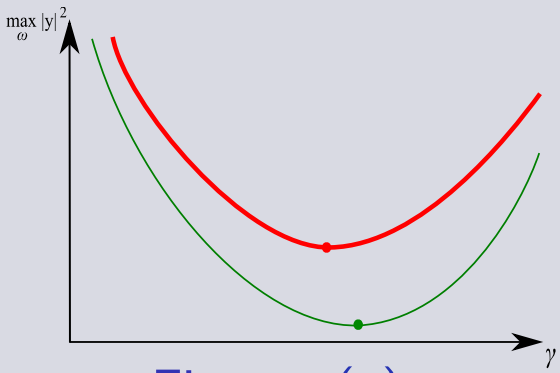


Figure: (a)

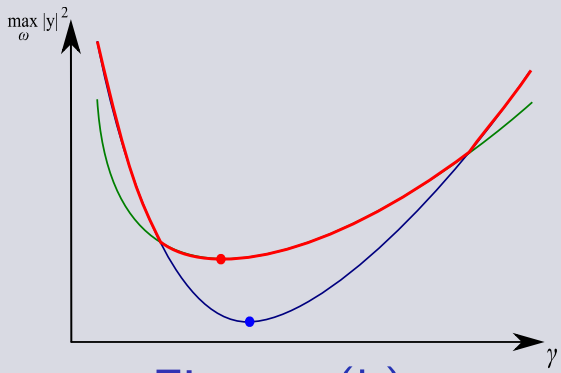


Figure: (b)

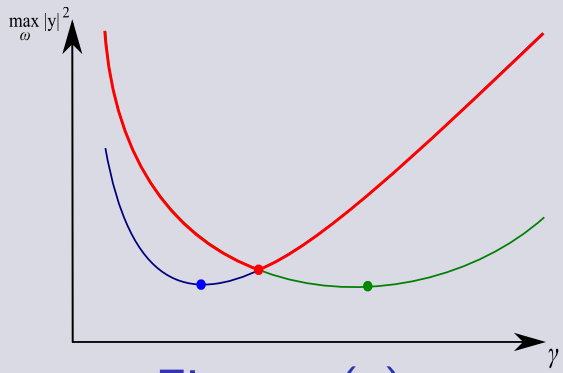


Figure: (c)

3 Possibilities:

(a) one dominant local maximum curve.

(b) no dominant curve, but global optimizer is a saddle point.

(c) global optimizer is not a saddle point.

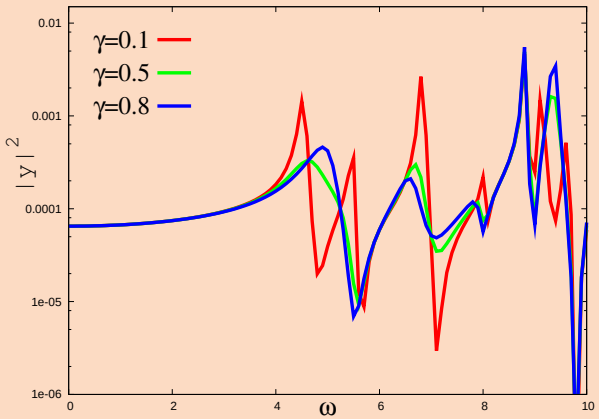
Minimax Optimization

Algorithm Scratch

Follow the global maximum curve, not the local maximum curves.

- Min Phase: Solve the original problem $\min_{\gamma} g(\gamma)$.
  - ✓ Max Phase: Solve the subproblem $g(\gamma_0) = \max_{\omega} |y(\omega, \gamma_0)|^2$, the global maximum for a given γ_0.

Some Details



Max Phase: Highly nonconvex, Many local optimizers.

Algorithm: Grid Search followed by Quasi Newton refinement.

Locates only one global optimizer.

Min Phase: Quasi Newton Method.

Can continue with the kink points without realizing it.

Algorithm: Stopping Condition

Either of the following two conditions is satisfied.

- $\left|\frac{\partial g}{\partial \gamma}\Big|_{\gamma=\gamma_k}\right| \leq \epsilon$: global optimizer is a saddle point.
- $|\gamma_k - \gamma_{k-1}| \leq \tau$: global optimizer is a kink point.

Note: All Quasi Newton steps uses backtracking strategy with Armijo condition.

Computational Cost of the Algorithm

For each step in Inner Max Phase

$$y = \ell^*(K + i\omega\gamma C - \omega^2 M)^{-1} f$$

$$\frac{\partial y}{\partial \omega} = \ell^*(K + i\omega\gamma C - \omega^2 M)^{-1}(2\omega M - i\gamma C)(K + i\omega\gamma C - \omega^2 M)^{-1} f$$

For each step in the Outer Min Phase

$$\frac{\partial y}{\partial \gamma} = \ell^*(K + i\omega\gamma C - \omega^2 M)^{-1}(-i\omega C)(K + i\omega\gamma C - \omega^2 M)^{-1} f$$

Minimax Optimization

Computational Cost of the Algorithm

- When the system order is large, computational costs of y, $\frac{\partial y}{\partial \omega}$ and $\frac{\partial y}{\partial \gamma}$ are very expensive!
- More time is spent on derivative computations than function evaluations.

Our Solution

Use Model Order Reduction (MOR) in optimization.

- It is well known that MOR can be used for function evaluations.
- We show MOR can also be used for derivative computations.

All data needed by the minimax algorithm can be computed with MOR, optimization time will be greatly reduced!

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

Basic Idea

Project a large scale linear system onto a Krylov subspace to get a reduced order linear system.

Krylov subspace

Krylov subspace:

$$\mathcal{K}_k\{A, b\} \triangleq \text{span}\{b, Ab, A^2b, \ldots, A^{k-1}b\}.$$

Generate it in a numerically stable way:

Arnoldi Process, Lanczos Method.

Arnoldi Process

$v_1 = b/\|b\|$
For $i = 2 \ldots k$
$w_i = Av_{i-1}$
Orthogonalize w_i to $v_1 \ldots v_{i-1}$, and normalize it to get v_i

Linear system with no design parameters

$$\begin{cases} (K - \alpha M)x = f \\ y = \ell^* x \end{cases} \qquad \textit{order } n \textit{ system}$$

MOR on this system

Let $V_k = [v_1, v_2, \ldots, v_k]$, $W_k = [w_1, w_2, \ldots, w_k] \in \mathbb{C}^{n \times k}$.
Approximate $x \approx V_k z \in \text{colspan}\{V_k\}$ and left multiply the first equation by W^*

$$\begin{cases} (\hat{K} - \alpha \hat{M})z = \hat{f} \\ \hat{y} = \hat{\ell}^* z \end{cases} \qquad \begin{array}{l} \text{order } k \text{ system} \\ k \ll n! \end{array}$$

where $\hat{K} = W^* K V$, $\hat{M} = W^* M V$, $\hat{f} = W^* f$, $\hat{\ell} = V^* \ell$.

Question: What about the accuracy?

KATHOLIEKE UNIVERSITEIT LEUVEN

Definition of Moment

Let $y = m_0 + m_1\alpha + m_2\alpha^2 + m_3\alpha^3 \ldots$

$\hat{y} = \hat{m}_0 + m_1\alpha + \hat{m}_2\alpha^2 + \hat{m}_3\alpha^3 \ldots$

m_i is called the i-th moment of y.

$\hat{m}_i$ is called the i-th moment of $\hat{y}$.

The idea behind Krylov Method

Match the first moments of y and $\hat{y}$ ($m_i = \hat{m}_i$),
so that for small αs,
$\hat{y}$ is a good approximation of y.

KATHOLIEKE UNIVERSITEIT LEUVEN

Accuracy of Krylov Method: Moment Matching Property

Theorem:

1. If $\text{colspan}\{V_k\} \supseteq$ the k-th right Krylov subspace $\mathcal{K}_k(K^{-1}M, K^{-1}b)$, the first k moments of y and $\hat{y}$ match;
2. If $\text{colspan}\{W_k\} \supseteq$ the k-th left Krylov subspace $\mathcal{K}_k(K^{-*}M^*, K^{-*}\ell^*)$, the first k moments of y and $\hat{y}$ match;
3. If $\text{colspan}\{V_k\} \supseteq$ the k-th right Krylov subspace and $\text{colspan}\{W_k\}$ spans the k-th left Krylov subspace $\mathcal{K}_k(K^{-*}M^*, K^{-*}\ell^*)$, the first $2k$ moments of y and $\hat{y}$ match.

Case 1 and 2 are called one-sided methods.

Case 3 is called two-sided method.

Challenge of Optimization on MOR

Recall the minimax optimization problem:

$$\min_{\gamma} \max_{\omega_L \le \omega \le \omega_H} |y(\omega,\gamma)|^2$$
$$\text{s.t.} \quad \begin{cases} y(\omega,\gamma) = \ell^* x(\omega,\gamma) \\ (K + i\omega\gamma C - \omega^2 M)x(\omega,\gamma) = f \end{cases}$$

It contains two parameters: frequency ω and design parameter γ.

How to use MOR for this system with 2 parameters?

Two Approaches

1. For each Max Phase, fix $\gamma = \gamma_0$ and do MOR with ω.
   - Representative: SOAR.
2. Parameterized MOR: Do MOR on both γ and ω.
   - Representative: PIMTAP.

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

Introduction

SOAR stands for Second Order ARnoldi Process.

Traditional Method: Use Arnoldi process on the linearized system.

Advantages of SOAR:

1. More accurate.
2. Structure preserving.

SOAR MOR

Project the system onto a "2nd Order Krylov Subspace".

$$\left\{\begin{array}{l} y = \ell^* x \\ (K + i\omega\gamma_0 C - \omega^2 M)x = f \end{array}\right. \xrightarrow{SOAR} \left\{\begin{array}{l} \hat{y} = \hat{\ell}^* z \\ (\hat{K} + i\omega\gamma_0 \hat{C} - \omega^2 \hat{M})z = \hat{f} \end{array}\right.$$

Order n System $\quad n \gg k \quad$ *Order k System*

$\hat{K} = W^* K V$, $\hat{C} = W^* C V$, $\hat{M} = W^* M V$, $\hat{f} = W^* f$ and $\hat{\ell} = V^* \ell$.

Moment Matching Property

Theorem:

1. If colspan$\{V_k\} \supseteq$ the k-th right 2nd order Krylov subspace, the first k moments of y and $\hat{y}$ match;
2. If colspan$\{W_k\} \supseteq$ the k-th left 2nd order Krylov subspace, the first k moments of y and $\hat{y}$ match;
3. If colspan$\{V_k\} \supseteq$ the k-th right 2nd order Krylov subspace and colspan$\{W_k\} \supseteq$ the k-th left 2nd order Krylov subspace, the first $2k$ moments of y and $\hat{y}$ match.

SOAR generates the 2-nd order Krylov subspaces in a numerically stable way.

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

PIMTAP MOR

PIMTAP: Parameterized Interconnect Macromodeling via a Two-directional Arnoldi Process.

- Could do MOR on more than one parameters.
- Structure preserving; relatively numerically stable.

Prototype:

$$\begin{cases} \big(G_0 + \gamma G_1 + s(C_0 + \gamma C_1)\big)x = b \\ y = \ell^* x \end{cases} \qquad \textit{Order } n \textit{ System}$$

$$\Downarrow \qquad n \gg k$$

$$\begin{cases} \big(\hat{G}_0 + \gamma \hat{G}_1 + s(\hat{C}_0 + \gamma \hat{C}_1)\big)z = \hat{b} \\ \hat{y} = \hat{\ell}^* z \end{cases} \qquad \textit{Order } k \textit{ System}$$

$$[\hat{G}_0, \hat{G}_1, \hat{C}_0, \hat{C}_1] = V^*[G_0, G_1, C_0, *C_1]V, \quad [\hat{b}, \hat{\ell}] = V^*[b, \ell].$$

V: generated by PIMTAP algorithm.

Multiparameter moments

Let $y = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} m_i^j s^i \gamma^j$,

m_i^j is called the (i, j)-th 2-parameter moment.

PIMTAP Algorithm

Generate the projection matrix V according to Moment Matching Pattern.

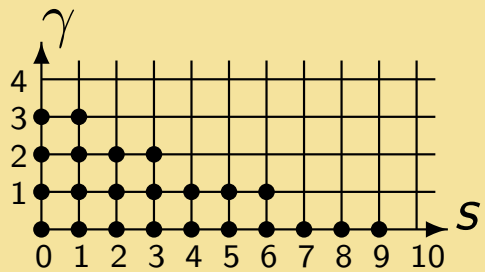


Figure: Example of Moment Matching Pattern.

All moments specified by the moment matching pattern will be matched for y and $\hat{y}$!

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

Using MOR in Minimax Optimization

Importance of Derivative Computations

- In the minimax optimization algorithm, more time is spent on derivative computations than function evaluations.
- If we can compute derivatives via the reduced model, the optimization time can further be saved.

SOAR case

SOAR prototype can be directly used for the Max Phase.
For derivative computations, we have the following theorem:

Theorem

If two-sided SOAR is used and the order of the reduced model is k:

- The first $2k-1$ moments of $\frac{\partial y}{\partial \omega}$ and $\frac{\partial \hat{y}}{\partial \omega}$ match.
- The first k moments of $\frac{\partial y}{\partial \gamma}$ and $\frac{\partial \hat{y}}{\partial \gamma}$ match.

No moment matching property for $\frac{\partial y}{\partial \gamma}$ and $\frac{\partial \hat{y}}{\partial \gamma}$ in one-sided SOAR!

Using MOR in Minimax Optimization

PIMTAP case

PIMTAP can be used in the Minimax Algorithm by $G_0 \leftarrow K$, $G_1 \leftarrow C$, $C_0 \leftarrow M$, $C_1 \leftarrow O$,$\gamma \leftarrow \gamma\omega$ and $s \leftarrow \omega^2$.

Theorem The (i,j)-th moments of ∇y and $\nabla \hat{y}$ match if the $(i+1,j)$-th and the $(i,j+1)$-th moments match for y and $\hat{y}$.

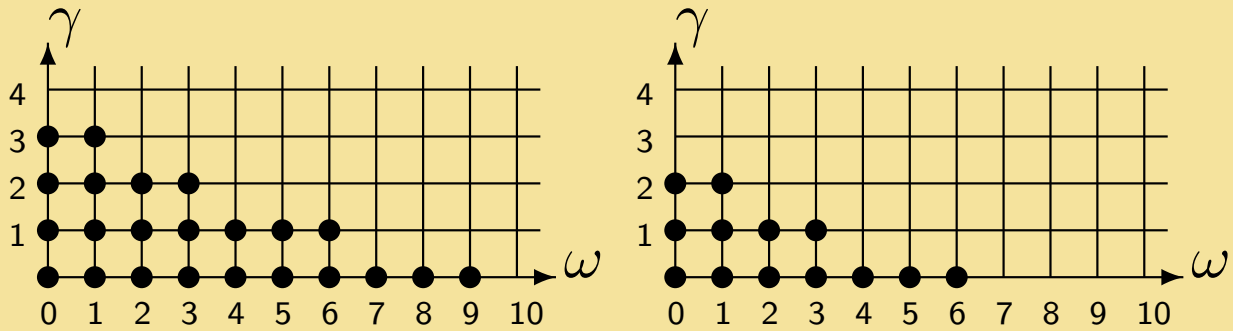


Figure: matched moments of y and ∇y in PIMTAP

Conclusion

Both two-sided SOAR and PIMTAP can compute the gradients with accuracy.

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

Numerical Results

Table: Example 1

	Direct method	SOAR	PIMTAP
Matrix size	15625	50	61
Optimizer computed	(10.1245, 0.2671)	(10.1245, 0.2671)	(10.1245, 0.2671)
CPU time	3508s	128s	22s

A relatively large example.
The optimizer is a saddle point.

Numerical Results

Table: Example 2

	Direct method	SOAR	PIMTAP
Matrix size	125	16	19
Optimizer computed	(8.8092, 0.9951)	(8.8092, 0.9951)	(8.8092, 0.9951)
CPU time	31s	1s	2s

A relatively small example.

The optimizer is a saddle point.

Numerical Results

Table: Example 3

	Direct method	SOAR	PIMTAP
Matrix size	216	35	44
Optimizer computed	(8.668, 0.2687)	(8.669, 0.2687)	(8.668, 0.2688)
CPU time	162s	27s	18s

A relatively small example.
The optimizer is a kink point.

KATHOLIEKE UNIVERSITEIT LEUVEN

Outline

1. Minimax Optimization
2. Introduction to Krylov based MOR
3. SOAR
4. PIMTAP
5. Using MOR in Minimax Optimization
6. Numerical Results
7. Conclusion

Conclusion

- MOR could drastically reduce the optimization time for large scale systems with high accuracy.
- Even for relatively small scale systems, MOR could also reduce the optimization time.
- In general, PIMTAP is more efficient than SOAR for large scale problems.
- The algorithm has superlinear convergence when one local maximum curve is dominant.
- When kink points exist in the global maximum curve, convergence rate will degrade, especially when the kink point is the optimizer, but convergence is always observed.

Thank you!