Model order reduction for PDE constrained optimization in vibrations

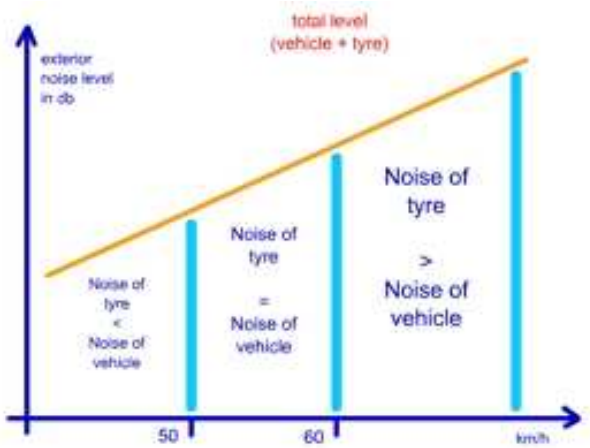
Karl Meerbergen
(Joint work with Yao Yue)

KU Leuven

SCEE12 – Zürich

Examples of vibrating systems

- Car tyres





- Windscreens



  - Structural damping
  - Choice of connection (glue) to the car

Examples of vibrating systems

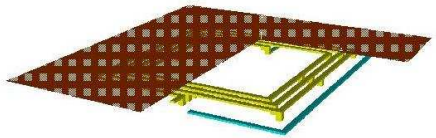
- Planes



- Bridge vibrating under footsteps and Thames wind



- Maxwell-equation – electrical circuits



- micro-gyroscope for navigation systems



Model order reduction (MOR) in optimization context

Dynamical system:

$$\begin{aligned} A(\omega,\gamma)x &= f \\ y &= d^T x \\ g &= \int_{\omega_{\min}}^{\omega_{\max}} |y|^2 d\omega \quad \text{(energy)} \end{aligned}$$



- Objective: minimize $g(\gamma)$
- Reduce the cost of computing y and ∇y:
  - Model reduction using moment matching = Padé approximation via Krylov methods
  - Allows for cheap computation of g and $\nabla_\gamma g$
- Assume uni-modal objective function

Outline

1. Motivation
2. MOR: Krylov-Padé methods
3. Parametric MOR
4. Trust region approach
5. Block Krylov method for low rank parameters
6. Conclusions

Left and right Krylov spaces

- For the linear case: $A(\omega) = I - \omega B$:
- Right Krylov space:

$$f, Bf, B^2 f, \ldots, B^{k-1} f$$

  Basis: $V_k = [v_1, \ldots, v_k]$
- Left Krylov space:

$$d, B^* d, B^{2*} d, \ldots, B^{(k-1)*} d$$

  Basis: $W_k = [w_1, \ldots, w_k]$
- Reduced model:

$$\begin{aligned} \widehat{A}(\omega)\widehat{x} &= \widehat{f} \\ \widehat{y} &= \widehat{d}^T \widehat{x} \end{aligned}$$

with $\widehat{A} = W_k{}^* A V_k$, $\widehat{f} = W_k{}^* f$, $\widehat{d} = V_k{}^* d$.

Moment matching

- Let $A(\omega) = I - \omega B$
- System output: $y = d^T (I - \omega B)^{-1} f$.

Theorem

Define the expansions:

$$\begin{aligned} y &= y_0 + \omega y_1 + \omega^2 y_2 + \cdots \\ \widehat{y} &= \widehat{y}_0 + \omega \widehat{y}_1 + \omega^2 \widehat{y}_2 + \cdots \end{aligned}$$

Let V_k and W_k be right and left Krylov basis resp., then

$$y_j = \widehat{y}_j \quad \textit{with} \quad j = 0, 1, \ldots, 2k-1 \ .$$

- Rational (Padé) approximation:

$$y(\omega) = \sum_{j=1}^{n} \frac{\rho_j}{\lambda_j - \omega} \qquad \widehat{y}(\omega) = \sum_{j=1}^{k} \frac{\widehat{\rho}_j}{\widehat{\lambda}_j - \omega}$$

- poles are Ritz values (approximate eigenvalues)

Gradient

- Build two-sided reduced model for fixed value of γ for
- $y = d^T A(\omega, \gamma_0)^{-1} f$
- Gradient:

$$\frac{dy}{d\gamma} = \left(A(\omega, \gamma_0)^{-*} d\right)^* \left(\frac{dA(\omega, \gamma_0)}{d\gamma}\right) \left(A(\omega, \gamma_0)^{-1} f\right)$$

- Blue part can be computed from right-Krylov space:

$$A(\omega, \gamma_0)^{-1} f \approx V_k(\widehat{A}^{-1}(\omega, \gamma_0)\widehat{f})$$

(k moments matched for $A(\omega, \gamma_0)^{-1} f$.)

- Red part can be computed from left-Krylov space:

$$A(\omega, \gamma_0)^{-*} d \approx W_k(\widehat{A}^{-*}(\omega, \gamma_0)\widehat{d})$$

(k moments matched for $A(\omega, \gamma_0)^{-*} d$.)

Gradient

- Full gradient

$$\frac{dy}{d\gamma} = \left(A(\omega,\gamma_0)^{-*}d\right)^* \left(\frac{dA(\omega,\gamma_0)}{d\gamma}\right) \left(A(\omega,\gamma_0)^{-1}f\right)$$

- Reduced gradient

$$\begin{aligned}\frac{d\widehat{y}}{d\gamma} &= \left(W_k\widehat{A}(\omega,\gamma_0)^{-*}\widehat{d}\right)^* \left(\frac{dA(\omega,\gamma_0)}{d\gamma}\right) \left(V_k\widehat{A}(\omega,\gamma_0)^{-1}\widehat{f}\right) \\ &= \left(\widehat{A}(\omega,\gamma_0)^{-*}\widehat{d}\right)^* \left(\frac{d\widehat{A}(\omega,\gamma_0)}{d\gamma}\right) \left(\widehat{A}(\omega,\gamma_0)^{-1}\widehat{f}\right)\end{aligned}$$

- $\frac{d\widehat{y}}{d\gamma}$ and $\frac{dy}{d\gamma}$ match k moments. [Antoulas, Beattie, Gugercin 2010] [Yue, M. 2011]
- Almost free computatation of the gradient, regardless the number of parameters!

Implementation issues

- Linear case: $A(\omega) = A_0 + \omega A_1$.

$$(A_0 + \omega A_1)x = f$$

- Replace by:

$$\begin{aligned} (I + \omega B)x &= b \quad \text{with} \quad B = A_0^{-1}A_1, b = A_0^{-1}f \\ y &= d^T x \end{aligned}$$

- Krylov space: sequence of matrix vector multiplies with B
- Computational cost:
  - Sparse matrix factorization of A_0
  - k sparse matrix vector products with A_1
  - k backward solves with A_0.

Nonlinear frequency dependence

$$(K + i\omega C - \omega^2 M)x = f$$

- 'Linearization':
  - Define matrices A and B

$$A = \begin{bmatrix} K & \\ & I \end{bmatrix} \qquad B = \begin{bmatrix} iC & -M \\ I & \end{bmatrix}$$

so that

$$(A - \omega B)\begin{pmatrix} x \\ \omega x \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}$$

This is called a linearization, a similar trick as the solution of second order ODE's.

Linearizations

- Higher order polynomials

$$(A_0 + \omega A_1 + \cdots + A_p \omega^p)x = f$$

- Methods based on Companion 'linearization' [Amiraslani, Corless, Lancaster, 2009]
- Transform to

$$\left(\begin{bmatrix} A_0 & A_1 & \cdots & A_{p-1} \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix} - \omega \begin{bmatrix} 0 & \cdots & \cdots & -A_p \\ I & & & 0 \\ & \ddots & & \vdots \\ & & I & 0 \end{bmatrix} \right) \begin{bmatrix} x \\ \omega x \\ \vdots \\ \omega^{p-1} x \end{bmatrix} = \begin{matrix} 0 \\ \\ \\ 0 \end{matrix}$$

- Can also be used for truly nonlinear problems (using Taylor expansion) [Van Beeumen, M., Michiels 2012]

Nonlinear output

- Quadratic output:

$$\begin{aligned}(A_0 - \omega A_1)x &= f \\ y &= x^* S x\end{aligned}$$

  with S a low rank matrix.

- reduced model:

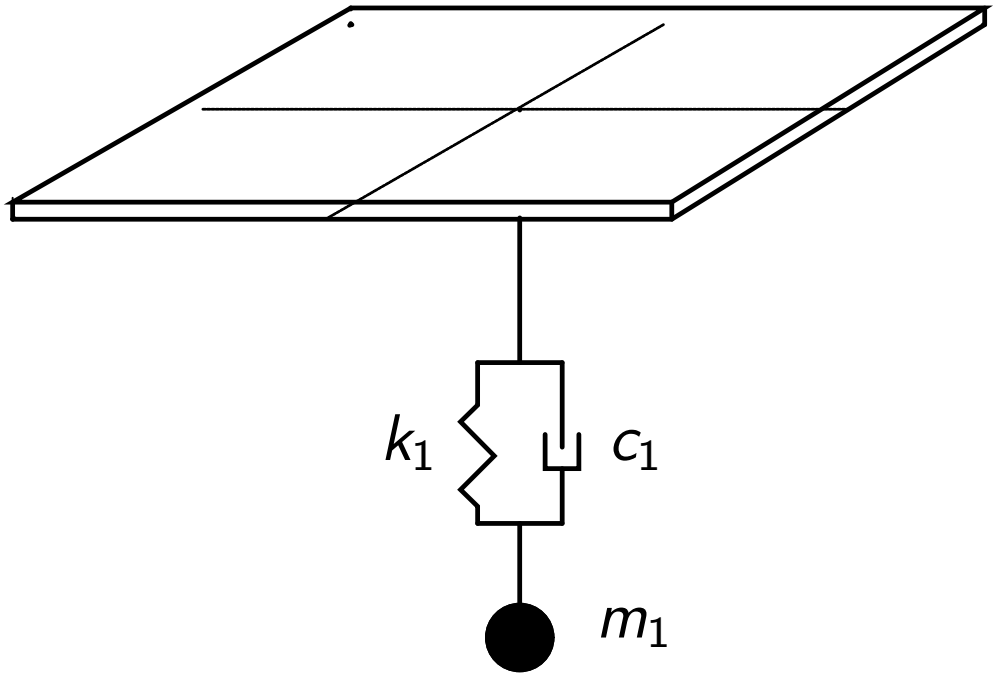
$$\begin{aligned}(\widehat{A}_0 - \omega \widehat{A}_1)\widehat{x} &= \widehat{f} \\ y &= \widehat{x}^* \widehat{S} \widehat{x}\end{aligned}$$

  with $\widehat{A}_0 = W^T A_0 V$, $\widehat{A}_1 = W^T A_1 V$, $\widehat{S} = V^T S V$ and $\widehat{f} = W^T f$.

- V: Krylov space with matrix $A_0^{-1} A_1$ and starting vector $A_0^{-1} f$
- W: Block-Krylov space with matrix $A_0^{-1} A_1$ and SV.
- Matching between $k + \sqrt{k}$ and $k + k$ moments
  [Van Beeumen, Van Nimmen, Lombaert, M., 2012]

Design optimization

- Determination of optimal parameters of a vibrating system
- Example: optimal parameters for a damper of a floor in a building near a noisy road



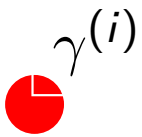
Design optimization

- Parametrized linear system:

$$\begin{cases} (K(\gamma) + i\omega C(\gamma) - \omega^2 M(\gamma))x &= f \\ y &= d^T x \end{cases}$$

- Find parameters γ so that
  - $\|y\|_2 = \int_0^{\omega_{\max}} |y|^2 d\omega$ is minimal
  - $\|y\|_\infty = \sup_0^{\omega_{\max}} |y|^2$ is minimal
- Assume: uni-modal, non necessarily smooth
- Expensive evaluation of y and the gradient

Overview



- Use MOR for function and gradient evaluation

Algorithm

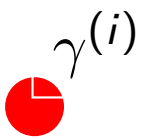
- We use the Damped BFGS optimization method
- On iteration i:
  - Build reduced model for $\gamma^{(i)}$ g
  - Compute g and ∇g from the reduced model
- Objective may not be smooth: use sufficient decrease condition and possibly backtracking after the BFGS step

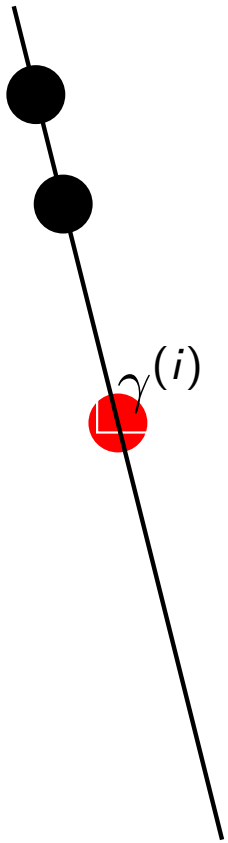
Numerical examples

- Floor with damper ($n = 29800$)
- Reduced model $k = 7$
- Determine optimal parameters c_1 and k_1

	Direct method	MOR
Matrix size	29800	7
Optimizer computed	$(12231609, 106031.18)$	$(12231614, 106031.22)$
Function value	$1.316093349\,10^{10}$	$1.316093349\,10^{10}$
CPU time	7626s	179s

Overview





- Use MOR for function and gradient evaluation
- Use PMOR for line search optimization + backtracking

Parametric MOR: PIMTAP

- Reduced model for equations of the following form:

$$\begin{aligned}(G_0 + \lambda G_1 + s(C_0 + \lambda C_1))x &= b \\ y &= d^* x\end{aligned}$$

- Perform moment matching, i.e. terms associated with $s^i \lambda^j$.
- Algorithms
  - Many, many papers, e.g. Lihong Feng.
  - PIMTAP [Li, Bai, Su, Zeng 2007], [Li, Bai, Su, Zeng 2008], [Li, Bai, Su 2009]
  - Subspace that contains the vectors:

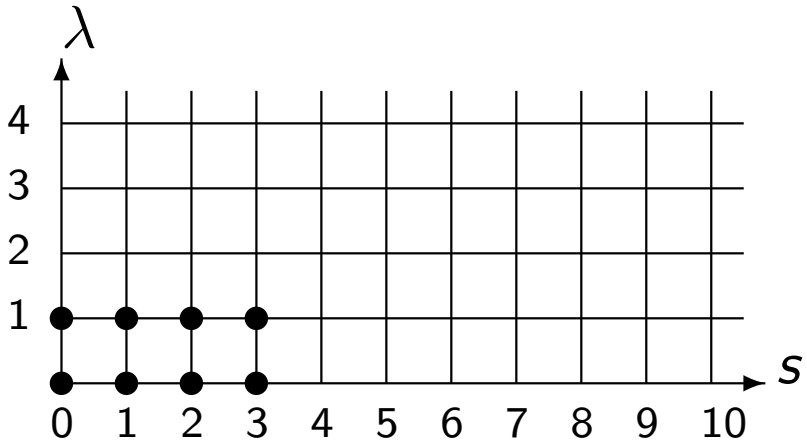
$$\begin{aligned}r_i^j &= 0 \quad i < 0, \quad \text{or} \quad j < 0 \\ r_0^0 &= G_0^{-1} b \\ r_i^j &= -G_0^{-1}(C_0 r_{i-1}^j + G_1 r_i^{j-1} + C_1 r_{i-1}^{j-1})\end{aligned}$$

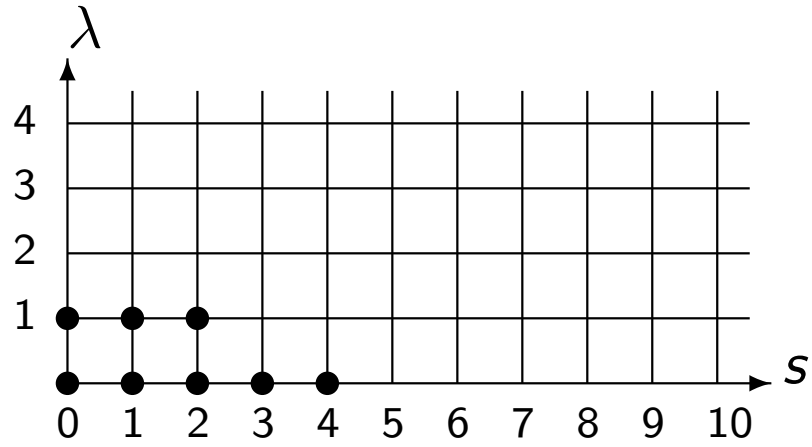
    These are the moments in the expansion $x \sim \sum r_i^j s^i \lambda^j$

- Two-sided PIMTAP
  - One PIMTAP for b and one for d.

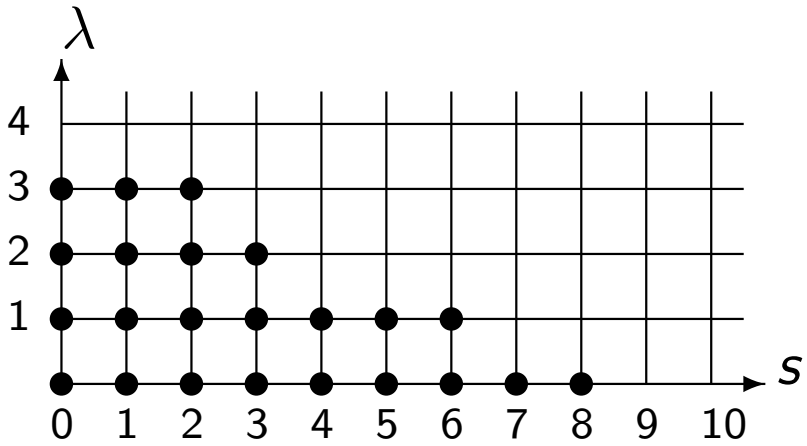
Moment matching of the gradient

- Moment matching patterns for left and right Krylov spaces



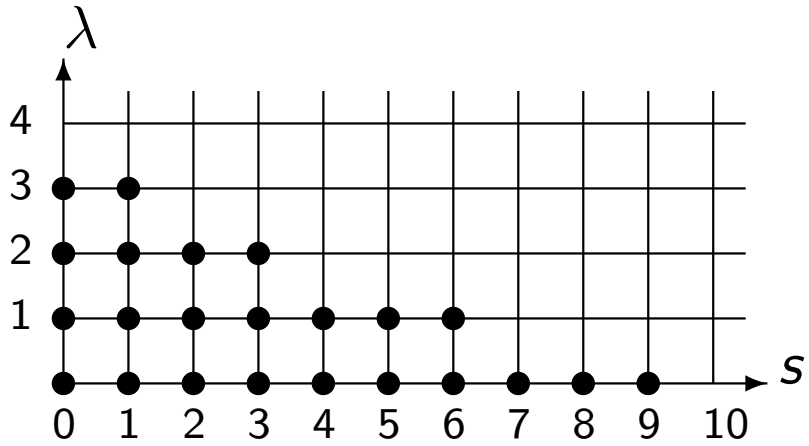


- Moment matching pattern for y

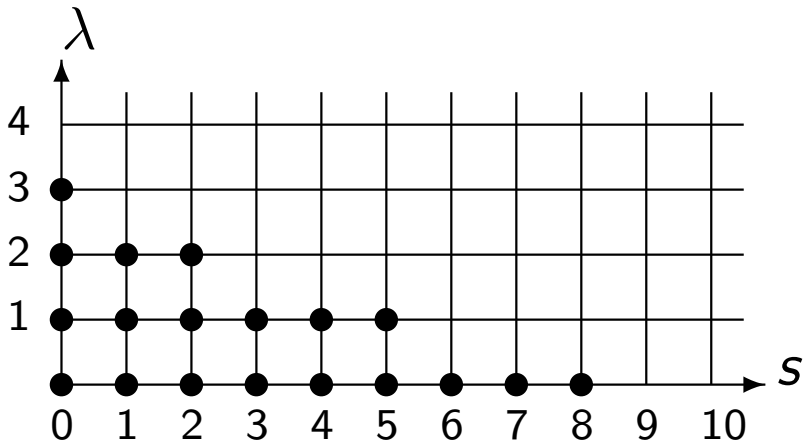


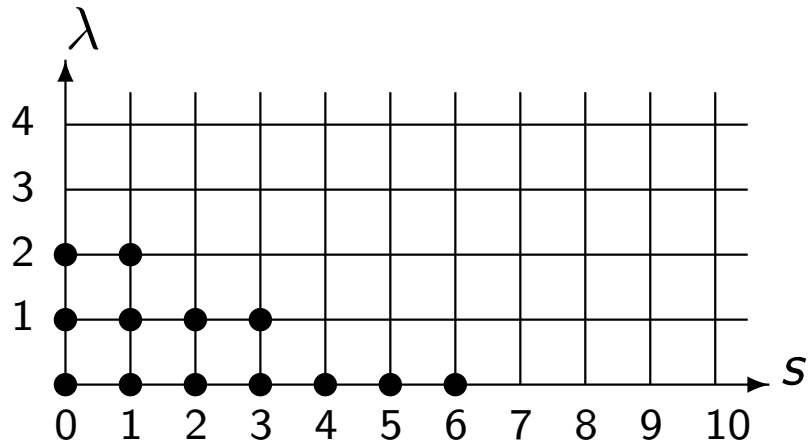
Moment matching of the gradient

- Moment matching pattern for y



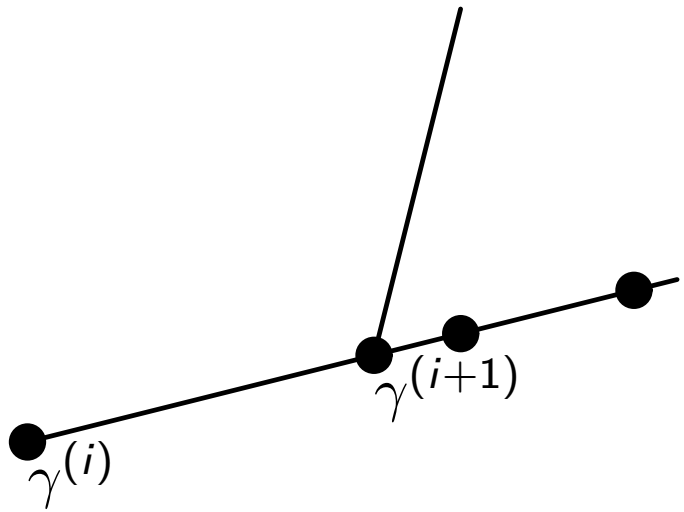
- Moment matching pattern for $\frac{\partial y}{\partial s}$ and $\frac{\partial y}{\partial \lambda}$





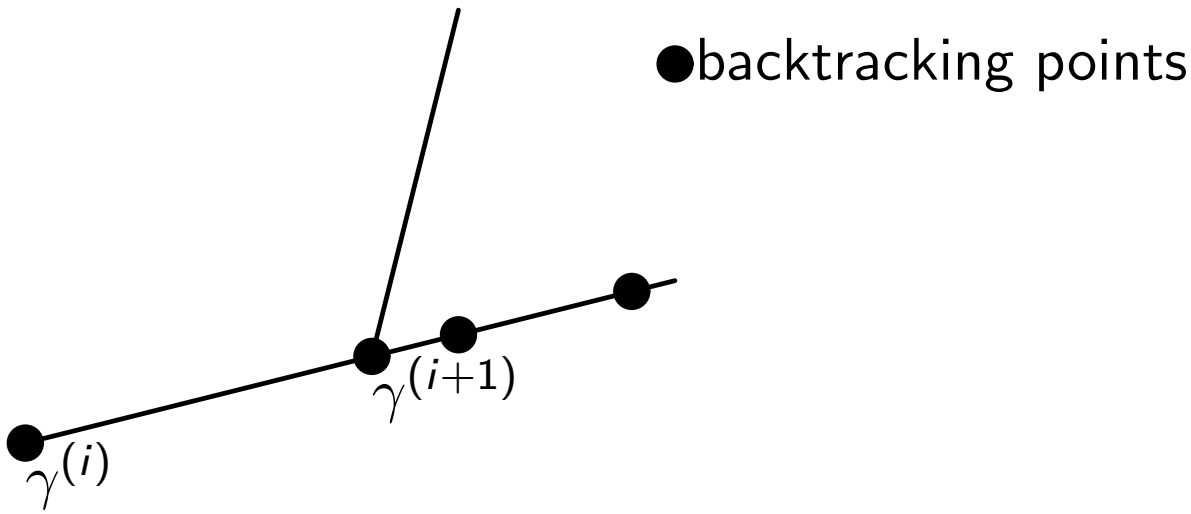
The MOR/PMOR Framework

- We can make a reduced model for ω and all parameters γ
- This is usually expensive and overkill: only reduce on the important directions
- So, we use PIMTAP for efficient line search:



  - The MOR Framework generates a reduced model for each γ accessed.
  - The PMOR Framework generates a reduced model for each line search iteration.

PMR framework

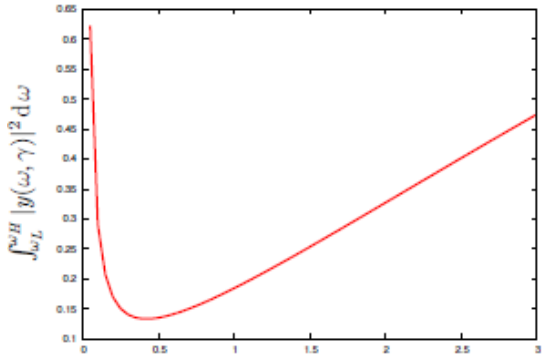


- For backtracking using Armijo line searcher:
  - the MOR Framework is better for smooth objective functions (The first trial point is often accepted for Quasi-Newton methods);
  - the PMOR Framework is better for non-smooth objective functions.
- For smooth objective functions, we can use exact line searches using the PMOR Framework.

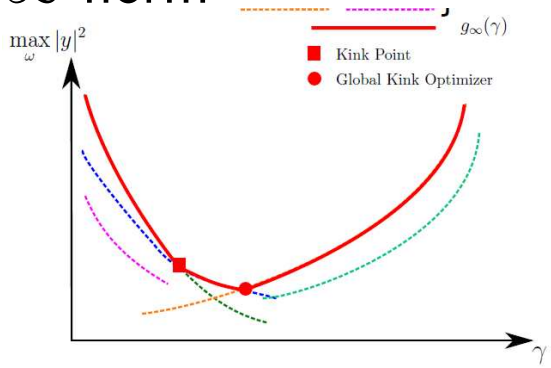
Example: min-max optimization

- Floor optimization problem (2 unknowns)
- Univariate objective function:

2 norm



∞ norm



- Comparison of MOR Framework and PMOR Framework

	2 norm			∞ norm		
	iter	k	Time	iter	k	Time
direct	11 (+6)	29,800	7626	15 (+108)	29,800	41,069
MOR	13 (+8)	7	179	15 (+120)	7	1,104
PMOR	12 (+4)	7+3	360	14 (+90)	7+3	417

Example: exact line search optimization

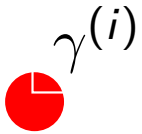
- Lamot bridge finite element model ($n = 25,962$)
- The goal is to determine the optimal stiffness and damping coefficient of four bridge dampers (=8 parameters).



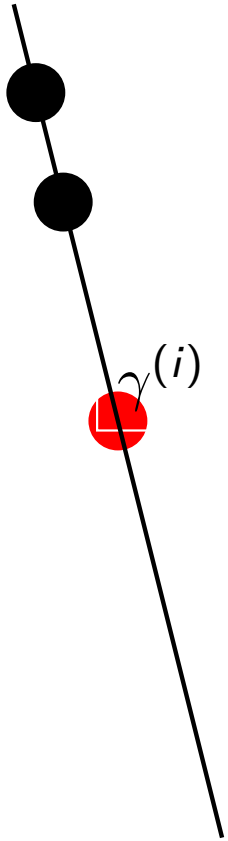
- Numerical results

	MOR	(backtracking) PMOR	Exact line search PMOR
order	12	11+4+2+1	11+4+2+1
iterations	70	73	27
time (s)	879	1830	703

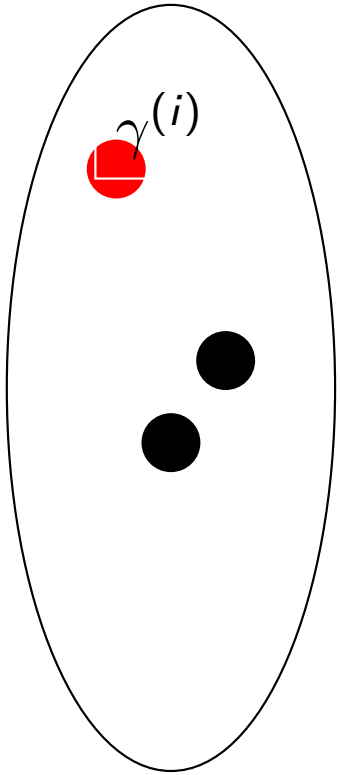
Overview



- Use MOR for function and gradient evaluation



- Use PMOR for line search optimization + backtracking



- Use interpolatory MOR for trust region optimization

Trust region approach

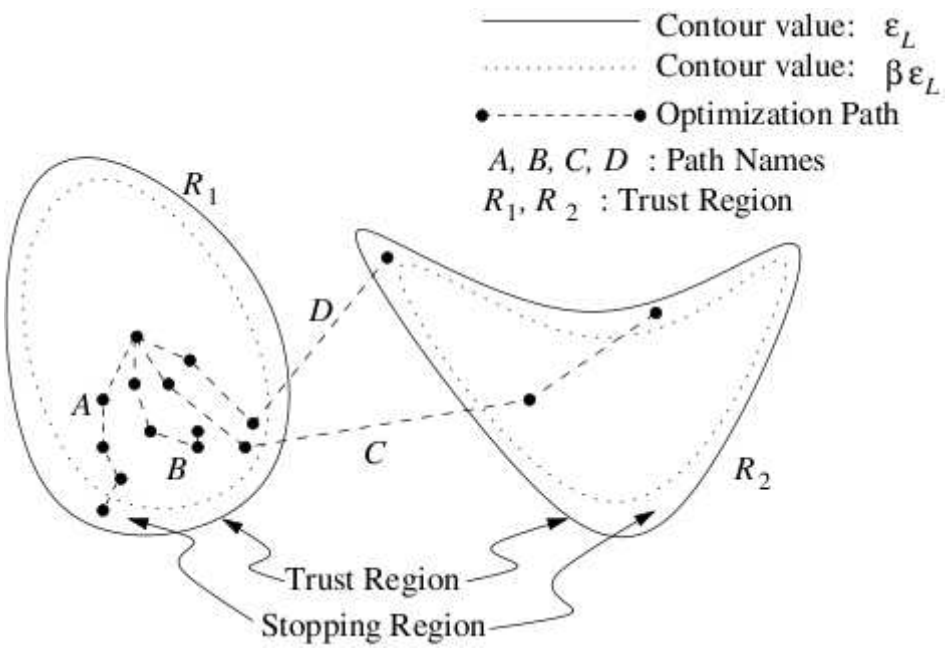
- The reduced model interpolates the exact function and the gradient in one point,
- so, we expect the model to be useful in a region around the interpolation point.
- On the ith iteration:
  1. Build Krylov space for $\gamma = \gamma^{(i)}$
  2. Use the subspace to make a reduced model in ω and γ:

$$\begin{aligned} \widehat{A}(\omega,\gamma)\widehat{x} &= \widehat{f} \\ \widehat{y} &= \widehat{d}^T\widehat{x} \end{aligned}$$

     with $\widehat{A} = W_k{}^* A V_k$, $\widehat{f} = W_k{}^* f$, $\widehat{d} = V_k{}^* d$.
  3. Find region where the error is acceptable (= trust region)
- Convergence test rewritten in terms of reduced models (that is what you have)
- Build new model only when required from convergence theory

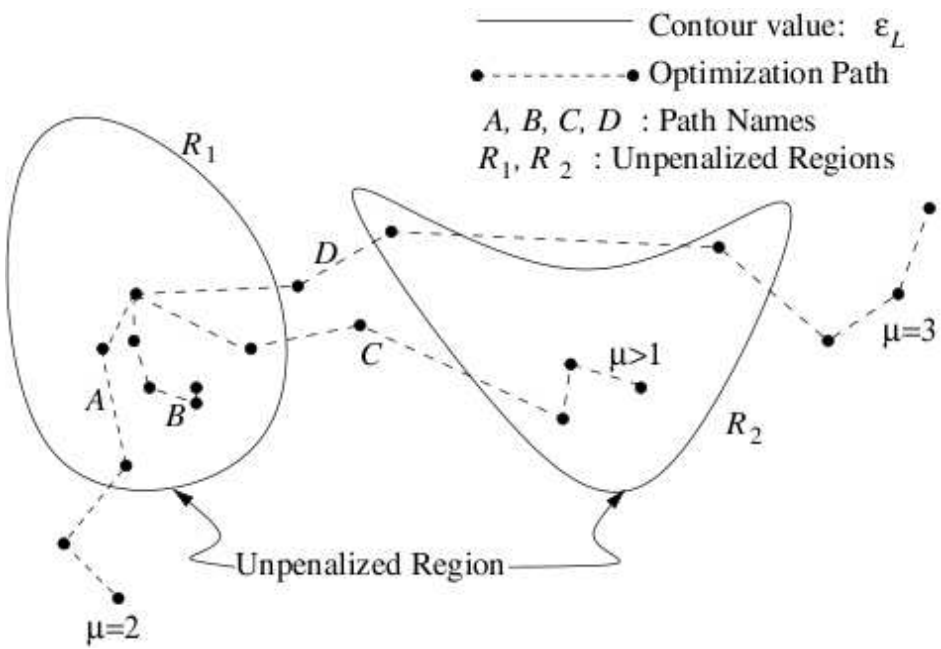
Subproblems

ETR



Terminate if close to boundary.

EP



Terminate if w is active for μ successive steps.

Example of the Lamot footbridge



- Lamot bridge finite element model ($n = 25,962$)
- The goal is to determine the optimal stiffness and damping coefficient of four bridge dampers (=8 parameters).

- Computation times:
  - without reduced modeling: 545 for one function evaluation, 70 function evaluations needed.
  - MOR Framework: 879 sec.
  - ETR: 200 sec.
  - EP: 189 sec.

Block Krylov method for low rank parameters

- Parametric system



$$\begin{aligned}(K - \omega^2 M + BC(\omega,\gamma)B^T)x &= f \\ y &= \varphi(x)\end{aligned}$$

where B is low rank r

- Reduced model:

$$\begin{aligned}(\widehat{K} - \omega^2 \widehat{M} + \widehat{B}C(\omega,\gamma)\widehat{B}^T)\widehat{x} &= \widehat{f} \\ y &= \varphi(V\widehat{x}) \\ \widehat{K} = V^T K V \qquad & \widehat{M} = V^T M V \\ \widehat{B} = V^T B \qquad & \widehat{f} = V^T f\end{aligned}$$

Block Krylov method for low rank parameters

- Reduced model:

$$
\begin{aligned}
(\widehat{K} - \omega^2 \widehat{M} + \widehat{B} C(\omega, \gamma) \widehat{B}^T)\widehat{x} &= \widehat{f} \\
y &= \varphi(V\widehat{x}) \\
\widehat{K} = V^T K V \qquad & \qquad \widehat{M} = V^T M V \\
\widehat{B} = V^T B \qquad & \qquad \widehat{f} = V^T f
\end{aligned}
$$

  with V Krylov basis for
  - Use block Krylov method:
  - starting block: $K^{-1}[f, B]$.
  - matrix: $K^{-1}M$
- Approximation properties:
  - For any γ, $2k$ moment match for ω
  - γ dependence is fully held by C and therefore represented exactly.
- We need only one reduced model for all optimization steps.

Example

2-norm Optimization of the Footbridge Damper Optimization Problem

	Nr Models	Order	Optimum	CPU Time
The MOR Framework	70	12	24.77751651	879s
The PMOR Framework	27	18	24.77751661	703s
ETR	3	20	24.78594112	205s
EP	3	20	24.7775166	190s
Block Arnoldi	1	30	24.77751815	20.4s

- Objective at the initial point: 142.34188.
- Damped-BFGS converges in 71 iterations.
- A single evaluation of g costs 540s.

Conclusions

- We proposed three type of methods to accelerate design optimization with (P)MOR.
- All of them are efficient for accelerating design optimization.
- Performance:

  Block Arnoldi $>$ *ETR/EP* $>$ *The (P)MOR Framework*

- Applicability:

  The MOR Framework, ETR/EP $>$ *The PMOR Framework, Block Arnoldi*

- Reliability:

  Depends on error estimation for the reduced model

Bibliography

1. K. Meerbergen. The solution of parametrized symmetric linear systems. *SIAM J. Matrix Anal. Appl.*, 24(4):1038–1059, 2003.
2. K. Meerbergen. Fast frequency response computation for Rayleigh damping. *International Journal of Numerical Methods in Engineering*, 73(1):96–106, 2008.
3. K. Meerbergen. The Quadratic Arnoldi method for the solution of the quadratic eigenvalue problem. *SIAM Journal on Matrix Analysis and Applications*, 30(4):1463–1482, 2008.
4. K. Meerbergen and Z. Bai. The Lanczos method for parameterized symmetric linear systems with multiple right-hand sides. *SIAM Journal on Matrix Analysis and Applications*, 31(4):1642–1662, 2010.
5. Wim Michiels, Elias Jarlebring, and Karl Meerbergen. Krylov based model order reduction of time-delay systems. *SIAM Journal on Matrix Analysis and Applications*, 32(4):1399–1421, 2011.
6. F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Review*, 43(2):235–286, 2001.
7. R. Van Beeumen, K. Meerbergen, and W. Michiels. A rational krylov method based on hermite interpolation for nonlinear eigenvalue problems. Technical Report TW613, KU Leuven, Department of Computer Science, Heverlee, Belgium, 2012.
8. R. Van Beeumen, K. Van Nimmen, G. Lombaert, and K. Meerbergen. Model reduction for dynamical systems with quadratic output. *International Journal of Numerical Methods in Engineering*, 91(3):229–248, 2012.
9. Y. Yue and K. Meerbergen. Using model order reduction for the design optimization of structures and vibrations. *International Journal of Numerical Methods in Engineering*, 2012.
10. Y. Yue and K. Meerbergen. Accelerating pde-constrained optimization by model order reduction with error control. Technical report TW611, Department of Computer Science, K.U.Leuven, Celestijnenenlaan 200A, 3001 Heverlee (Leuven), Belgium, 2012.