Petrov-Galerkin Approximations in (Multi-)Linear Algebra from the Lens of Optimization

SIAM Conference on Optimzation
University of Endinburgh

Nicolas Venkovic, Aryaman Jeendgar and Hartwig Anzt

Chair of Computational Mathematics
School of Computation, Information and Technology
Technical University of Munich

June 2, 2026

Outline I

Introduction

The "What?", the "Why?" and the "How?"

- A few *approximation problems* of interest to us:
  - **Linear solves** with $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$:

$$\text{Find } x \in \mathbb{R}^n \text{ s.t. } Ax \approx b.$$

  - **Inversion** of $A \in \mathbb{R}^{n\times n}$:

$$\text{Find } M \in \mathbb{R}^{n\times n} \text{ s.t. } M \approx A^{-1}.$$

  - **Low-rank CP representation** of $\mathfrak{X} \in \mathbb{R}^{n_1\times\cdots\times n_d}$:

$$\text{Find } (U_1, \dots, U_d) \in \mathbb{R}^{n_1\times r} \times \cdots \times \mathbb{R}^{n_d\times r} \text{ s.t. } [\![U_1, \dots, U_d]\!] \approx \mathfrak{X}$$

  for some $r \ll \min\{n_1, \dots, n_d\}$ where

$$[\![U_1, \dots, U_d]\!] := \sum_{k=1}^{r} \underbrace{U_1[:,k]}_{u_k^{(1)} \in \mathbb{R}^{n_1}} \circ \cdots \circ \underbrace{U_d[:,k]}_{u_k^{(d)} \in \mathbb{R}^{n_d}} .$$

- Interests in both:
  - unconstrained variants of these problems, and the case of
  - **structured** (e.g., sparse, nonnegative, symmetric, definite, ...) **iterates**.

Best approximation in Hilbert spaces

- From approximation theory:

  For each element $x \in \mathcal{V}$ of a **Hilbert space** $\mathcal{V}$,
  and given convex **approximation "space"** $\mathcal{K} \subset \mathcal{V}$ closed in $\mathcal{V}$,
  there exists a unique **best approximation** $y \in \mathcal{K}$ of x in $\mathcal{K}$:

  $$\exists\,!\, y \in \mathcal{K} \text{ s.t. } \|x - y\| < \|x - z\| \;\forall\, z \in \mathcal{K} \setminus \{y\}.$$

  That is, the **forward approximation error norm** induced by the inner product admits a unique **global minimizer** $y \in \mathcal{K}$.

  Moreover, the **error** of the best approximation $y \in \mathcal{K}$ is **orthogonal** to the **search space** $\mathcal{K}$:

  $$x - y \perp \mathcal{K}.$$

  We say that y is the **orthogonal projection** of x onto $\mathcal{K}$.

  When the search space $\mathcal{K}$ is **finite-dimensional**, finding such orthogonal projections hopefully becomes a **tractable algebraic problem**.

Petrov-Galerkin approximation in Hilbert spaces

- When $x \in \mathcal{V}$ is **unknown**, we are left seeking zero(s) of residual maps

$$r : \mathcal{V} \to \mathcal{W}.$$

For example, in unconstrained problems, we seek to zero the following residuals:

- Linear solves: $$\begin{aligned} r : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto b - Ax \end{aligned}$$
- Matrix inverting: $$\begin{aligned} R : \mathbb{R}^{n\times n} &\to \mathbb{R}^{n\times n} \\ M &\mapsto I_n - AM \end{aligned}$$
- Low-rank CP representation: $$\begin{aligned} \mathfrak{R} : \mathbb{R}^{n_1\times r} \times \cdots \times \mathbb{R}^{n_d\times r} &\to \mathbb{R}^{n_1\times\cdots\times n_d} \\ (U_1, \ldots, U_d) &\mapsto \mathfrak{X} - [\![U_1, \ldots, U_d]\!] \end{aligned}.$$

- In this context, **Petrov-Galerkin approximations** are defined by

$$\boxed{\text{Find } y \in \mathcal{K} \subset \mathcal{V} \text{ s.t. } r(y) \perp \mathcal{L} \subset \mathcal{W}}$$

where $\mathcal{K}$ is an approximation **search space**, and $\mathcal{L}$ is a **space** of **orthogonality constraints**.

Approximate linear solves

Relation to stationary conditions of quadratic models

- If $A \in \mathbb{R}^{n \times n}$ is invertible then,
  for every $b \in \mathbb{R}^n$, there exists a unique vector $x \in \mathbb{R}^n$ such that $Ax = b$.
  This vector, $x = A^{-1}b$, is the unique stationary point of every quadratic model

  $$m : x \in \mathbb{R}^n \mapsto m(x) := \frac{x^T A x}{2} - x^T b + c \text{ with } c \in \mathbb{R}.$$

  It is

  - a unique minimizer (resp., maximizer) iff $A \succ 0_{n \times n}$ (resp., $A \prec 0_{n \times n}$),
  - a unique saddle point iff A is indefinite.

State-of-the-art iterations

Essential to the deployment of (Petrov-)Galerkin methods is the construction of orthogonal bases for the search space. There are

- **Short-recurrence** iterations:
  - Conjugate gradient (CG):
    CG iterates arise from a Galerkin condition on a recursively updated one-dimensional affine search space:

$$x_{i+1} \in x_i + \mathsf{span}\{p_i\} \;\; \text{such that} \;\; r_{i+1} := b - Ax_{i+1} \perp \mathsf{span}\{p_i\}$$

    with a Petrov-Galerkin (i.e., A-orthogonality) condition on a recursively updated one-dimensional affine space for the search direction:

$$p_i \in r_i + \mathsf{span}\{p_{i-1}\} \;\; \text{such that} \;\; p_i \perp A\,\mathsf{span}\{p_{i-1}\}.$$

State-of-the-art iterations

Essential to the deployment of (Petrov-)Galerkin methods is the construction of orthogonal bases for the search space. There are

- **Short-recurrence** iterations:
  - Conjugate gradient (CG):
    Remarkably, the recursive nature of these one-dimensional spaces renders CG iterates equivalently defined as a Galerkin method on a $(i+1)$-dimensional Krylov search space:

$$x_{i+1} \in x_0 + \mathcal{K}_{i+1}(A, r_0) \text{ such that } r_{i+1} \perp \mathcal{K}_{i+1}(A, r_0)$$

    where $\mathcal{K}_{i+1}(A, r_0) := \text{span}\{r_0, Ar_0, \ldots, A^i r_0\}$ denotes the vectors formed by application to r_0 of polynomials of A of degree $j \leq i$.

State-of-the-art iterations

Essential to the deployment of (Petrov-)Galerkin methods is the construction of orthogonal bases for the search space. There are

- **Short-recurrence** iterations:
  - Conjugate gradient (CG):
    If A is symmetric positive definite, then the A-norm of the forward error is minimized by the CG iterate:

$$\|x_{i+1} - x\|_A = \min_{y \in x_0 + \mathcal{K}_{i+1}(A, r_0)} \|y - x\|_A.$$

$\implies$ CG iterates do *not* minimize the residual norm $\|r_{i+1}\|$ over the search space.

State-of-the-art iterations

Essential to the deployment of (Petrov-)Galerkin methods is the construction of orthogonal bases for the search space. There are

- **Short-recurrence** iterations:
  - Faber and Manteuffel:
    Give necessary conditions for the construction of orthogonal bases by short-recurrence iteration.

    $\Longrightarrow$ For non-normal matrices, the existence of short-recurrence formulas generally require applications of A^T, e.g., as in BiCGStab.

State-of-the-art iterations

Essential to the deployment of (Petrov-)Galerkin methods is the construction of orthogonal bases for the search space. There are

- **Full-orthogonalization**-based iterations:
  - General minimal residual (GMRES):
    For general matrices, the standard method of choice is GMRES.
    A GMRES iterate is formed by Petrov-Galerkin projection in a Krylov subspace:

$$x_{i+1} \in x_0 + \mathcal{K}_{i+1}(A, r_0) \text{ such that } r_{i+1} \perp A\mathcal{K}_{i+1}(A, r_0)$$

    thereby minimizing the residual norm $\|r_{i+1}\|$.
    Every new iterate requires orthogonality be enforced to all previous spaces.

    $\Longrightarrow$ Resort to GMRES prohibited in cases of strict limitations over memory and data movement.

Approximate matrix inversion

Relation to stationary conditions of quadratic models

- If A is invertible then,
  then there exists a unique matrix $M \in \mathbb{R}^{n \times n}$ such that $AM = I_n$.

  This matrix, $M = A^{-1}$, is the unique stationary point of every quadratic model

  $$m : M \in \mathbb{R}^{n\times n} \mapsto m(M) := \frac{(M, AM)_F}{2} - \text{tr}(M) + c \ \text{ with } \ c \in \mathbb{R}.$$

  As before, it is

  - a unique minimizer (resp., maximizer) iff $A \succ 0_{n\times n}$ (resp., $A \prec 0_{n\times n}$),
  - a unique saddle point iff A is indefinite.

- Practical limitations arise due to the iterates' sizes in $O(n^2)$:
  - Short-recurrence iteration only,
  - Sparse mode implementations for when A is sparse.

Approximate inverses by conjugate gradient iteration

- Conjugate gradient (CG):
  CG is not commonly used in practice to compute approximate inverses, but we advocate for its sparse-mode implementation for the computation of SPAIs in Venkovic and Anzt (2025).

  A CG iterate of approximate inverse of an SPD matrix A is given by

  $$M_{i+1} \in M_i + \mathsf{span}\{P_i\} \text{ such that } R_{i+1} := I_n - AM_{i+1} \perp \mathsf{span}\{P_i\}$$

  with an update of the search direction by

  $$P_i \in R_i + \mathsf{span}\{P_{i-1}\} \text{ such that } P_i \perp A\,\mathsf{span}\{P_{i-1}\}.$$

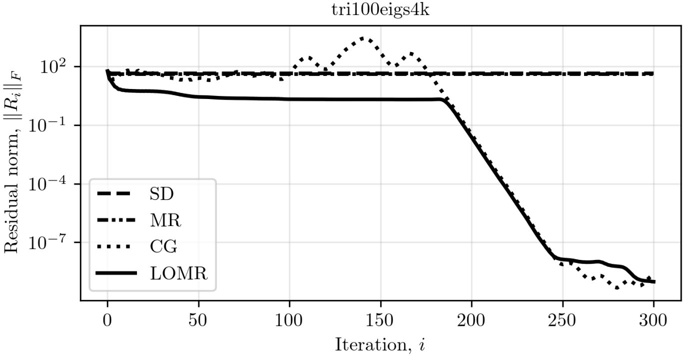
  As remarkably as for vector iterates, this is equivalently stated by

  $$M_{i+1} \in M_0 + \mathcal{K}_{i+1}(A, R_0) \text{ such that } R_{i+1} \perp \mathcal{K}_{i+1}(A, R_0).$$

N. Venkovic and H. Anzt (2025) Global iterative methods for sparse approximate inverses of symmetric positive-definite matrices. arXiv preprint, arXiv:2511.09753.

Approximate inverses by conjugate gradient iteration

- Conjugate gradient (CG):
  CG minimizes the A-norm of the forward error, i.e., $\|A^{-1} - M\|_A$.
  In practice, the CG residual iterate norm $\|R_i\|$ can show large variations.
  Efficiently preconditioning this iteration is difficult.



Every iteration can be very costly.

N. Venkovic and H. Anzt (2025) Global iterative methods for sparse approximate inverses of symmetric positive-definite matrices. arXiv preprint, arXiv:2511.09753.

Approximate inverses by minimal residual method

- Minimum residual (MR):
  Alternatively, MR performs a one-dimensional Petrov-Galerkin recursive update of the approximate inverse:

$$M_{i+1} \in M_i + \mathsf{span}\{R_i\} \text{ s.t. } R_{i+1} := I_n - AM_{i+1} \perp A\,\mathsf{span}\{R_i\}$$

  which achieves a monotone decrease of residual norm as

$$\|R_{i+1}\| := \min_{M \in M_i + \mathsf{span}\{R_i\}} \|I_n - AM\|.$$

  Although MR converges quadratically in sufficient proximity of A^{-1}, it is still too slow for some initial iterates.

Chow, E., & Saad, Y. (1998). Approximate inverse preconditioners via sparse-sparse iterations. SIAM Journal on Scientific Computing, 19(3), 995-1023.

Approximate inverses by locally optimal minimal residual method

- Locally optimal minimal residual (LOMR):
  LOMR locally enriches the search and constraints spaces of MR iterations:

$$\begin{aligned} \text{Find } \ & M_{i+1} \in M_i + \text{span}\{R_i, P_{i-1}\} \\ \text{such that } \ & R_{i+1} := I_n - AM_{i+1} \perp A\,\text{span}\{R_i, P_{i-1}\} \end{aligned}$$

  which is deployed by solving a quadratic scalar equation, and so that

$$\text{LOMR's } \ \|R_i\| \leq \text{MR's } \ \|R_i\| \text{ for } i = 1, 2, \ldots.$$

  In practice, LOMR sometimes significantly outperforms MR and CG.

N. Venkovic and H. Anzt (2025) Global iterative methods for sparse approximate inverses of symmetric positive-definite matrices. arXiv preprint, arXiv:2511.09753.

Structured approximate inverses

Of particular interest is the case where structural constraints are applied on the approximate inverse. E.g., sparsity, definiteness, symmetry, ...

- Sparse approximate inverses (SPAI):

  For every prescribed support supp(M) of nonzero values of an SPAI, there is an associated vector space from which one can readily extract a unique best approximate inverse.
  More generally, for any prescribed value $k < n^2$ of nonzero components, there exists a unique **best k-sparse approximate inverse**.
  But finding the sparsity pattern of that best k-sparse approximate inverse is challenging. Two approaches exist:

  - Incorporating dropping strategies, that is, nonlinear maps

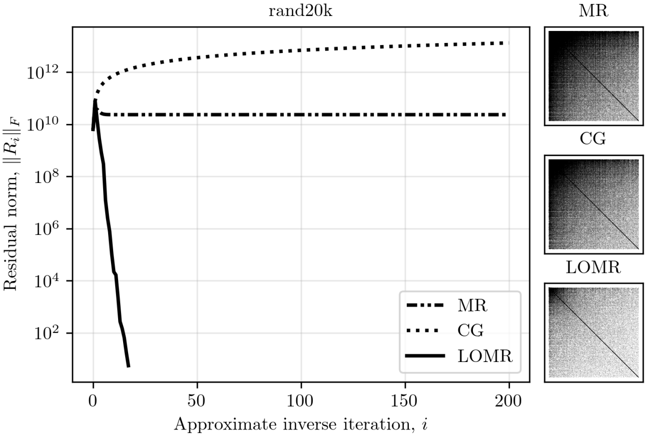
    $$T_k : \mathbb{R}^{n\times n} \to \{X \in \mathbb{R}^{n\times n}, |\mathrm{supp}(X)| \leq k\}$$

    of which k-sparse approximations are fixed points.
    E.g., hard thresholding and more sophisticated schemes.
  - Regularization.

Experimental results

- k-sparse approximate inverse iterates are formed by applying dropping strategies to every new iterate.

  Here are sample results for the rand20k matrix:



The matrix can be found at `github.com/venkovic/matrix-market`.

Approximate low-rank CP tensor representation

Low-rank CP tensor approximation

- To construct low-rank CP tensor approximations of the d-way array $\mathfrak{X} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$, we make use of the gradients

$$\nabla_{U_\ell} g : \mathbb{R}^{n_1 \times r} \times \cdots \times \mathbb{R}^{n_d \times r} \to \mathbb{R}^{n_\ell \times r} \text{ for } \ell = 1, \ldots, d$$

  of the non-convex objective function

$$g : (U_1, \ldots, U_d) \mapsto \|\mathfrak{X} - [\![U_1, \ldots, U_d]\!]\|^2.$$

- A Petrov-Galerkin gradient descent is formulated with approximates

$$U_{\ell,i+1} \in U_{\ell,i} + \mathsf{span}\left\{\nabla_{U_\ell} g(U_{1,i}, \ldots, U_{d,i})\right\}$$

  where $r \cdot n_\ell$ orthogonality constraints need be applied to close the system for $\ell = 1, \ldots, d$.
  For matrices ($d = 2$), we proceed with the following constraints

$$R_{i+1} V_i \perp \mathsf{span}\{R_i V_i\}$$
$$R_{i+1}^T U_i \perp \mathsf{span}\{R_i^T U_i\}$$

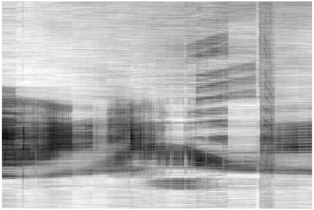
  which generalize to d-way arrays.

Comparison of methods

- Similarly as for approximate inverses, the Petrov-Galerkin gradient descent can be made locally optimal for faster convergence.

  Low-rank approximations achieved in 5 seconds:

  Petrov-Galerkin gradient descent

  Locally optimal variant

  For $d = 2$, all minima of the objective function are global,
  and every similarity transformation introduces an equally good low-rank approximation.

  This complicates the case of structured (e.g., k-sparse) approximations, where low values of the objective function may be achieved independently of whether the iterate assumes the structure wanted.

Comparison of methods

As d increases, the value of the objective function becomes more indicative of whether the structure of data is captured by an iterate, but the landscape of the objective function also changes.

Closing remarks

Closing remarks

- Take-away message:
  - Petrov-Galerkin methods allow for significant flexibility in seeking approximate solutions to (non)linear equations over a wide variety of Hilbert spaces.
- Dissemination:
  - Find this presentation at venkovic.github.io/research.
  - Preprint: N. Venkovic and H. Anzt (2025) Global iterative methods for sparse approximate inverses of symmetric positive-definite matrices. arXiv preprint, arXiv:2511.09753.
- Related ongoing work:
  - Mixed precision Petrov-Galerkin methods.
  - Parallel computation of sparse approximate inverses.
  - Structured low-rank matrix approximation.
  - Structured data recovery.
- Next talk:
  - Structured iterative approximations in numerical (multi-)linear algebra.

August 13 @ JuliaCon 2026.