Sparse Days in Saint-Girons IV, June 2022

Towards efficient randomized limited memory preconditioners for variational data assimilation

Alexandre Scotto Di Perrotolo

Youssef Diouane Selime Gürol Xavier Vasseur

Université de Toulouse, ISAE-SUPAERO, Toulouse, France

Outline

Outline

The weighted nonlinear least-squares problem

We consider the problem of fitting n state variables of a dynamical system to $m \ll n$ noisy observations, given a noisy prior state estimate, which can be formalized as

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} \|\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}, t)\|^2_{R^{-1}} + \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{x_0}\|^2_{B^{-1}} \,,$$

where $\mathcal{H}$ is the prediction operator, $B \in \mathbb{R}^{n \times n}$ the a priori state error covariance matrix and $R \in \mathbb{R}^{m \times m}$ the observation error covariance matrix.

The weighted nonlinear least-squares problem

We consider the problem of fitting n state variables of a dynamical system to $m \ll n$ noisy observations, given a noisy prior state estimate, which can be formalized as

$$\min_{x\in\mathbb{R}^n} f(x) = \frac{1}{2}\|\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}, t)\|^2_{R^{-1}} + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x_0}\|^2_{B^{-1}},$$

where $\mathcal{H}$ is the prediction operator, $B \in \mathbb{R}^{n\times n}$ the a priori state error covariance matrix and $R \in \mathbb{R}^{m\times m}$ the observation error covariance matrix.

Given a current approximate solution x_j and approximating $d_j = y - \mathcal{H}(x_j, t) \approx y - H_j x_j$, the solution using the Gauss-Newton method (*Nocedal et al., 2006*) computes $x_{j+1} = x_j + s_j$ with the j-th descent direction s_j satisfying

$$\underbrace{\left(B^{-1} + H_j^\mathsf{T} R^{-1} H_j\right)}_{=A_j} s_j = \underbrace{B^{-1}(x_c - x_j) + H_j^\mathsf{T} R^{-1} d_j}_{=b_j}$$

The weighted nonlinear least-squares problem

We consider the problem of fitting n state variables of a dynamical system to $m \ll n$ noisy observations, given a noisy prior state estimate, which can be formalized as

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}\|\boldsymbol{y} - \mathcal{H}(\boldsymbol{x}, t)\|^2_{R^{-1}} + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x_0}\|^2_{B^{-1}},$$

where $\mathcal{H}$ is the prediction operator, $B \in \mathbb{R}^{n \times n}$ the a priori state error covariance matrix and $R \in \mathbb{R}^{m \times m}$ the observation error covariance matrix.

Given a current approximate solution x_j and approximating $d_j = y - \mathcal{H}(x_j, t) \approx y - H_j x_j$, the solution using the Gauss-Newton method (*Nocedal et al., 2006*) computes $x_{j+1} = x_j + s_j$ with the j-th descent direction s_j satisfying

$$\underbrace{\left(B^{-1} + H_j^\mathsf{T} R^{-1} H_j\right)}_{= A_j} s_j = \underbrace{B^{-1}(x_c - x_j) + H_j^\mathsf{T} R^{-1} d_j}_{= b_j}$$

Solving the linearized subproblem

The descent directions s_j are computed using an iterative method with B as a right preconditioner. If $s_j = B\bar{s}_j$, then $\bar{s}_j$ is such that

$$\bar{A}_j \bar{s}_j = b_j, \quad \text{with} \quad \begin{cases} \bar{A}_j = I_n + H_j^\intercal R^{-1} H_j B & \text{(new system matrix)}, \\ B\bar{A}_j = \bar{A}_j^\intercal B & \text{(B-symmetry)}. \end{cases}$$

Solving the linearized subproblem

The descent directions s_j are computed using an iterative method with B as a right preconditioner. If $s_j = B\bar{s}_j$, then $\bar{s}_j$ is such that

$$\bar{A}_j \bar{s}_j = b_j, \quad \text{with} \quad \begin{cases} \bar{A}_j = I_n + H_j^\intercal R^{-1} H_j B & \text{(new system matrix)}, \\ B\bar{A}_j = \bar{A}_j^\intercal B & \text{(B-symmetry)}. \end{cases}$$

▷ Conjugate gradient in the B inner product (*Gürol, 2013*) with preconditioner $\bar{P}_j$.

Solving the linearized subproblem

The descent directions s_j are computed using an iterative method with B as a right preconditioner. If $s_j = B\bar{s}_j$, then $\bar{s}_j$ is such that

$$\bar{A}_j \bar{s}_j = b_j, \quad \text{with} \quad \begin{cases} \bar{A}_j = I_n + H_j^\top R^{-1} H_j B & \text{(new system matrix)}, \\ B\bar{A}_j = \bar{A}_j^\top B & (B\text{-symmetry}). \end{cases}$$

▷ Conjugate gradient in the B inner product (*Gürol, 2013*) with preconditioner $\bar{P}_j$.

To enhance the solution process of the sequence, it is common to update the preconditioner, starting from $\bar{P}_1 = I_n$. Then for all $j \geq 2$, we consider

$$\bar{P}_j = \bar{P}_{j-1}$$

Solving the linearized subproblem

The descent directions s_j are computed using an iterative method with B as a right preconditioner. If $s_j = B\bar{s}_j$, then $\bar{s}_j$ is such that

$$\bar{A}_j \bar{s}_j = b_j, \quad \text{with} \quad \begin{cases} \bar{A}_j = I_n + H_j^\top R^{-1} H_j B & \text{(new system matrix)}, \\ B\bar{A}_j = \bar{A}_j^\top B & (B\text{-symmetry}). \end{cases}$$

▷ Conjugate gradient in the B inner product (*Gürol, 2013*) with preconditioner $\bar{P}_j$.

To enhance the solution process of the sequence, it is common to update the preconditioner, starting from $\bar{P}_1 = I_n$. Then for all $j \geq 2$, we consider

$$\bar{P}_j = \bar{P}_{j-1} + V_j \left(\Lambda_j^{-1} - I_k \right) V_j^\top B,$$

where $V_j \in \mathbb{R}^{n \times k}$ and $\Lambda_j \in \mathbb{R}^{k \times k}$ contain *exact* eigenvectors and eigenvalues of $\bar{P}_{j-1}\bar{A}_j$ respectively.

Solving the linearized subproblem

The descent directions s_j are computed using an iterative method with B as a right preconditioner. If $s_j = B\bar{s}_j$, then $\bar{s}_j$ is such that

$$\bar{A}_j \bar{s}_j = b_j, \quad \text{with} \quad \begin{cases} \bar{A}_j = I_n + H_j^\top R^{-1} H_j B & \text{(new system matrix)}, \\ B\bar{A}_j = \bar{A}_j^\top B & \text{(B-symmetry)}. \end{cases}$$

⊳ Conjugate gradient in the B inner product (*Gürol, 2013*) with preconditioner $\bar{P}_j$.

To enhance the solution process of the sequence, it is common to update the preconditioner, starting from $\bar{P}_1 = I_n$. Then for all $j \geq 2$, we consider

$$\bar{P}_j = \bar{P}_{j-1} + V_j \left(\Lambda_j^{-1} - I_k \right) V_j^\top B,$$

where $V_j \in \mathbb{R}^{n \times k}$ and $\Lambda_j \in \mathbb{R}^{k \times k}$ contain approximate eigenvectors and eigenvalues of $\bar{P}_{j-1}\bar{A}_j$ respectively, computed using randomized methods.

Outline

Randomized methods to compute eigenpairs

Let $\bar{A}, \bar{P}$ be two B-symmetric matrices, we are interested in finding pairs (λ, v) such that,

$$\bar{P}\bar{A}\, v = \lambda\, v$$

Randomized methods to compute eigenpairs

Let $\bar{A}, \bar{P}$ be two B-symmetric matrices, we are interested in finding pairs (λ, v) such that,

$$\bar{P}\bar{A}\,v = \lambda\,v \iff \underbrace{B\bar{A}\,v = \lambda\,B\bar{P}^{-1}\,v}_{\text{Generalized symmetric eigenvalue problem}}\,.$$

Randomized methods to compute eigenpairs

Let $\bar{A}, \bar{P}$ be two B-symmetric matrices, we are interested in finding pairs (λ, v) such that,

$$\bar{P}\bar{A}\,v = \lambda\,v \iff \underbrace{B\bar{A}\,v = \lambda\,B\bar{P}^{-1}\,v}_{\text{Generalized symmetric eigenvalue problem}} .$$

Prior algorithms: (Saibaba, Lee and Kitanidis, 2016) for generalized Hermitian eig. problem.

Randomized methods to compute eigenpairs

Let $\bar{A}, \bar{P}$ be two B-symmetric matrices, we are interested in finding pairs (λ, v) such that,

$$\bar{P}\bar{A}\,v = \lambda\,v \iff \underbrace{B\bar{A}\,v = \lambda\,B\bar{P}^{-1}\,v}_{\text{Generalized symmetric eigenvalue problem}}.$$

Prior algorithms: (Saibaba, Lee and Kitanidis, 2016) for generalized Hermitian eig. problem.

1. Need to apply $B\bar{A}$ and $(B\bar{P}^{-1})^{-1} = \bar{P}B^{-1}$.
2. Require to perform a QR factorization in the $B\bar{P}^{-1}$ inner product.

Randomized methods to compute eigenpairs

Let $\bar{A}, \bar{P}$ be two B-symmetric matrices, we are interested in finding pairs (λ, v) such that,

$$\bar{P}\bar{A}\,v = \lambda\,v \iff \underbrace{B\bar{A}\,v = \lambda\,B\bar{P}^{-1}\,v}_{\text{Generalized symmetric eigenvalue problem}}.$$

Prior algorithms: (Saibaba, Lee and Kitanidis, 2016) for generalized Hermitian eig. problem.

1. Need to apply $B\bar{A}$ and $(B\bar{P}^{-1})^{-1} = \bar{P}B^{-1}$.
2. Require to perform a QR factorization in the $B\bar{P}^{-1}$ inner product.

Let $\bar{A}, \bar{P}$ be two B-symmetric matrices, we are interested in finding pairs (λ, v) such that,

$$\bar{P}\bar{A}\,v = \lambda\,v \iff \underbrace{B\bar{A}\,v = \lambda\,B\bar{P}^{-1}\,v}_{\text{Generalized symmetric eigenvalue problem}} .$$

Prior algorithms: (Saibaba, Lee and Kitanidis, 2016) for generalized Hermitian eig. problem.

1. Need to apply $B\bar{A}$ and $(B\bar{P}^{-1})^{-1} = \bar{P}B^{-1}$.
2. Require to perform a QR factorization in the $B\bar{P}^{-1}$ inner product.

Our randomized algorithms

Design of a dedicated class of randomized methods following two classical steps:

Our randomized algorithms

Design of a dedicated class of randomized methods following two classical steps:

- Step 1: randomized subspace iteration (*Halko et al., 2011*)

  Construct search space $V_q = (\bar{P}\bar{A})^q\,\Omega$ with random matrix $\Omega \in \mathbb{R}^{n\times p}$ and $q \geq 1$.

- Step 2: Rayleigh-Ritz method in the B inner product

  Extract $k \leq p$ eigenpairs from V_q by solving the projected eigenvalue problem

$$B\bar{A}v \quad = \lambda\, B\bar{P}^{-1}v \quad , \quad v \in \mathbb{R}^n.$$

Our randomized algorithms

Design of a dedicated class of randomized methods following two classical steps:

- Step 1: randomized subspace iteration (*Halko et al., 2011*)

  Construct search space $V_q = (\bar{P}\bar{A})^q\,\Omega$ with random matrix $\Omega \in \mathbb{R}^{n\times p}$ and $q \geq 1$.

- Step 2: Rayleigh-Ritz method in the B inner product

  Extract $k \leq p$ eigenpairs from V_q by solving the projected eigenvalue problem

$$B\bar{A}\,v \quad = \lambda\,B\bar{P}^{-1}\,v \quad , \quad v \in \mathbb{R}^n.$$

Our randomized algorithms

Design of a dedicated class of randomized methods following two classical steps:

- Step 1: randomized subspace iteration (*Halko et al., 2011*)

  Construct search space $V_q = (\bar{P}\bar{A})^q \Omega$ with random matrix $\Omega \in \mathbb{R}^{n \times p}$ and $q \geq 1$.

- Step 2: Rayleigh-Ritz method in the B inner product

  Extract $k \leq p$ eigenpairs from V_q by solving the projected eigenvalue problem

$$V_q^\top B \bar{A} V_q y = \lambda V_q^\top B \bar{P}^{-1} V_q y, \quad y \in \mathbb{R}^p.$$

Our randomized algorithms

Design of a dedicated class of randomized methods following two classical steps:

- Step 1: randomized subspace iteration (*Halko et al., 2011*)

  Construct search space $V_q = (\bar{P}\bar{A})^q \, \Omega$ with random matrix $\Omega \in \mathbb{R}^{n \times p}$ and $q \geq 1$.

- Step 2: Rayleigh-Ritz method in the B inner product

  Extract $k \leq p$ eigenpairs from V_q by solving the projected eigenvalue problem

$$V_q^\top B \bar{A} \, V_q \, y = \lambda \, V_q^\top B \bar{P}^{-1} \, V_q \, y, \quad y \in \mathbb{R}^p.$$

▷ Existing approach: Construct V_q such that $V_q^\top B \bar{P}^{-1} V_q = I_p$

$+$ Solve a standard symmetric eig. problem.

Our randomized algorithms

Design of a dedicated class of randomized methods following two classical steps:

- Step 1: randomized subspace iteration (*Halko et al., 2011*)

  Construct search space $V_q = (\bar{P}\bar{A})^q\,\Omega$ with random matrix $\Omega \in \mathbb{R}^{n\times p}$ and $q \geq 1$.

- Step 2: Rayleigh-Ritz method in the B inner product

  Extract $k \leq p$ eigenpairs from V_q by solving the projected eigenvalue problem

$$V_q^{\mathsf{T}} B \bar{A}\, V_q\, y = \lambda\, V_q^{\mathsf{T}} B \bar{P}^{-1}\, V_q\, y, \quad y \in \mathbb{R}^p.$$

▷ Existing approach: Construct V_q such that $V_q^{\mathsf{T}} B \bar{P}^{-1} V_q = I_p$
$+$ Solve a standard symmetric eig. problem.

▷ Our approach: Construct V_q such that $V_q^{\mathsf{T}} V_q = I_p$
$+$ Solve a generalized symmetric eig. problem.

Framework of the theoretical analysis

Let us consider the eigenvalue decomposition

$$\bar{P}\bar{A}V = V\Lambda, \quad \text{with } V \text{ a } B\bar{P}^{-1}\text{-orthogonal matrix.}$$

Framework of the theoretical analysis

Let us consider the eigenvalue decomposition

$$\bar{P}\bar{A}V = V\Lambda, \quad \text{with } V \text{ a } B\bar{P}^{-1}\text{-orthogonal matrix.}$$

Let $V_k \in \mathbb{R}^{n\times k}$ denote the matrix containing the k dominant eigenvectors, then the best rank k approximation of $\bar{P}\bar{A}$ is

$$(\bar{P}\bar{A})_k = V_k V_k^\intercal B\bar{A} = \pi_{B\bar{P}^{-1}}(V_k)\,\bar{P}\bar{A},$$

where $\pi_{B\bar{P}^{-1}}(V_k)$ denotes the $B\bar{P}^{-1}$-orthogonal projector onto $\mathcal{R}(V_k)$.

Framework of the theoretical analysis

Let us consider the eigenvalue decomposition

$$\bar{P}\bar{A}V = V\Lambda, \quad \text{with } V \text{ a } B\bar{P}^{-1}\text{-orthogonal matrix.}$$

Let $V_k \in \mathbb{R}^{n\times k}$ denote the matrix containing the k dominant eigenvectors, then the best rank k approximation of $\bar{P}\bar{A}$ is

$$(\bar{P}\bar{A})_k = V_k V_k^{\mathsf{T}} B\bar{A} = \pi_{B\bar{P}^{-1}}\left(V_k\right)\bar{P}\bar{A},$$

where $\pi_{B\bar{P}^{-1}}\left(V_k\right)$ denotes the $B\bar{P}^{-1}$-orthogonal projector onto $\mathcal{R}(V_k)$.

▷ Analysis: Let $Z \in \mathbb{R}^{n\times p}$ be such that $\mathcal{R}(Z) \approx \mathcal{R}(V_k)$, then we consider the low rank approximation $\pi_{B\bar{P}^{-1}}\left(Z\right)\bar{P}\bar{A}$

Let us consider the eigenvalue decomposition

$$\bar{P}\bar{A}V = V\Lambda, \quad \text{with } V \text{ a } B\bar{P}^{-1}\text{-orthogonal matrix.}$$

Let $V_k \in \mathbb{R}^{n\times k}$ denote the matrix containing the k dominant eigenvectors, then the best rank k approximation of $\bar{P}\bar{A}$ is

$$(\bar{P}\bar{A})_k = V_k V_k^\mathsf{T} B\bar{A} = \pi_{B\bar{P}^{-1}}(V_k)\,\bar{P}\bar{A},$$

where $\pi_{B\bar{P}^{-1}}(V_k)$ denotes the $B\bar{P}^{-1}$-orthogonal projector onto $\mathcal{R}(V_k)$.

⊳ Analysis: Let $Z \in \mathbb{R}^{n\times p}$ be such that $\mathcal{R}(Z) \approx \mathcal{R}(V_k)$, then we consider the low rank approximation $\pi_{B\bar{P}^{-1}}(Z)\,\bar{P}\bar{A}$ and its associated error in weighted spectral norm $|||\cdot|||_2$:

$$|||\pi_{B\bar{P}^{-1}}(Z)\ \bar{P}\bar{A} - \bar{P}\bar{A}|||_2$$

Framework of the theoretical analysis

Let us consider the eigenvalue decomposition

$$\bar{P}\bar{A}V = V\Lambda, \quad \text{with } V \text{ a } B\bar{P}^{-1}\text{-orthogonal matrix.}$$

Let $V_k \in \mathbb{R}^{n\times k}$ denote the matrix containing the k dominant eigenvectors, then the best rank k approximation of $\bar{P}\bar{A}$ is

$$(\bar{P}\bar{A})_k = V_k V_k^{\mathsf{T}} B\bar{A} = \pi_{B\bar{P}^{-1}}(V_k)\,\bar{P}\bar{A},$$

where $\pi_{B\bar{P}^{-1}}(V_k)$ denotes the $B\bar{P}^{-1}$-orthogonal projector onto $\mathcal{R}(V_k)$.

⊳ Analysis: Let $Z \in \mathbb{R}^{n\times p}$ be such that $\mathcal{R}(Z) \approx \mathcal{R}(V_k)$, then we consider the low rank approximation $\pi_{B\bar{P}^{-1}}(Z)\,\bar{P}\bar{A}$ and its associated error in weighted spectral norm $|||\cdot|||_2$:

$$|||\pi_{B\bar{P}^{-1}}(Z)\,(\bar{P}\bar{A})_k - (\bar{P}\bar{A})_k|||_2 \quad \text{is ideally equal to } 0.$$

Average-case analysis of low rank approximation error

Theorem (*Y.D., S.G., A. Scotto Di Perrotolo, X.V., 2022*)

Let $Z = (\bar{P}\bar{A})^q\,\Omega$ with $q \geq 1$ and $\Omega \in \mathbb{R}^{n\times p} \sim \mathcal{N}(0, I_n)$, and let us denote by $\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues of $\bar{P}\bar{A}$. Then for all $1 \leq k \leq p-2$ one has

$$\mathbb{E}\left[\ |||[I_n - \pi_{B\bar{P}^{-1}}(Z)](\bar{P}\bar{A})_k|||_2\ \right] \qquad ,$$

Average-case analysis of low rank approximation error

Theorem (*Y.D., S.G., A. Scotto Di Perrotolo, X.V., 2022*)

Let $Z = (\bar{P}\bar{A})^q\,\Omega$ with $q \geq 1$ and $\Omega \in \mathbb{R}^{n\times p} \sim \mathcal{N}(0, I_n)$, and let us denote by $\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues of $\bar{P}\bar{A}$. Then for all $1 \leq k \leq p-2$ one has

$$\mathbb{E}\left[\, \| [I_n - \pi_{B\bar{P}^{-1}}(Z)](\bar{P}\bar{A})_k \|_2 \right] \leq c_2 + O\left(\frac{1}{\sqrt{p}}\right) \quad ,$$

Average-case analysis of low rank approximation error

Theorem (*Y.D., S.G., A. Scotto Di Perrotolo, X.V., 2022*)

Let $Z = (\bar{P}\bar{A})^q \Omega$ with $q \geq 1$ and $\Omega \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, I_n)$, and let us denote by $\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues of $\bar{P}\bar{A}$. Then for all $1 \leq k \leq p-2$ one has

$$\mathbb{E}\left[|||[I_n - \pi_{B\bar{P}^{-1}}(Z)](\bar{P}\bar{A})_k|||_2 \right] \leq c_2 + O\left(\frac{1}{\sqrt{p}}\right) = O\left(\left(\frac{\lambda_{k+1}}{\lambda_k}\right)^q\right),$$

Average-case analysis of low rank approximation error

Theorem (*Y.D., S.G., A. Scotto Di Perrotolo, X.V., 2022*)

Let $Z = (\bar{P}\bar{A})^q \Omega$ with $q \geq 1$ and $\Omega \in \mathbb{R}^{n \times p} \sim \mathcal{N}(0, I_n)$, and let us denote by $\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues of $\bar{P}\bar{A}$. Then for all $1 \leq k \leq p-2$ one has

$$\mathbb{E}\left[\left\| [I_n - \pi_{B\bar{P}^{-1}}(Z)](\bar{P}\bar{A})_k \right\|_2 \right] \leq c_2 + O\left(\frac{1}{\sqrt{p}}\right) = O\left(\left(\frac{\lambda_{k+1}}{\lambda_k}\right)^q\right),$$

Additional results:

- A similar result holds in weighted Frobenius norm.
- Our analysis integrates the case of non-standard Gaussian matrices.

Average-case analysis of low rank approximation error

Theorem (*Y.D., S.G., A. Scotto Di Perrotolo, X.V., 2022*)

Let $Z = (\bar{P}\bar{A})^q\,\Omega$ with $q \geq 1$ and $\Omega \in \mathbb{R}^{n\times p} \sim \mathcal{N}(0, I_n)$, and let us denote by $\lambda_1 \geq \cdots \geq \lambda_n$ the eigenvalues of $\bar{P}\bar{A}$. Then for all $1 \leq k \leq p-2$ one has

$$\mathbb{E}\left[\, \| [I_n - \pi_{B\bar{P}^{-1}}(Z)](\bar{P}\bar{A})_k \|_2 \right] \leq c_2 + O\left(\frac{1}{\sqrt{p}}\right) = O\left(\left(\frac{\lambda_{k+1}}{\lambda_k}\right)^q\right),$$

Additional results:

- A similar result holds in weighted Frobenius norm.
- Our analysis integrates the case of non-standard Gaussian matrices. In particular, if we denote by $\mathbf{Cov}(\Omega)$ the covariance matrix of Ω then one has

$$\mathbf{Cov}(\Omega) = B\bar{P}^{-1} \implies c_2 = 0.$$

Outline

The 4D-Var data assimilation test problem

We propose an application of the proposed method to a 4D-Var data assimilation problem.

The Lorenz 95 model: State vector $x = (X_1, \ldots, X_n)$ whose components satisfy

$$\frac{dX_l}{dt} = -X_{l-2}X_{l-1} + X_{l-1}X_{l+1} - X_l + F, \quad 1 \leq l \leq n$$

with periodic boundary conditions. We set $n = 500$ state variables and $N = 24$ time steps implying operators of size up to 12500×12500.

Three scenarios: Sensitivity to the number of observations

- `LowObs`: 120 observations are made ($\approx 1\%$ of observations),
- `MedObs`: 1260 observations are made ($\approx 10\%$ of observations),
- `HighObs`: 2520 observations are made ($\approx 20\%$ of observations).

Details of the experiments

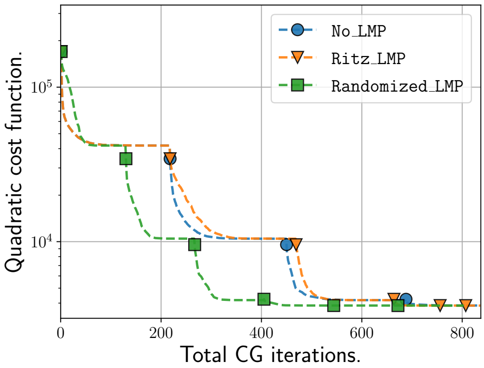
General setting:

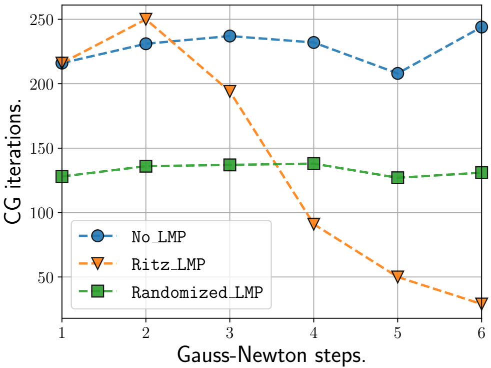
- We perform 6 Gauss-Newton steps.
- Tolerance for the CG convergence is set to $\varepsilon = 10^{-4}$ with a maximum of 250 iterations.
- Randomized algorithms compute $k = 30$ eigenpairs with $p = 50$ samples and $q = 1$.

Practical preconditioning strategies:

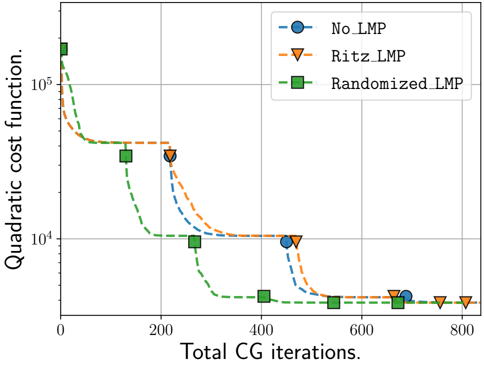
- `No_LMP`: the preconditioner is not updated.
- `Ritz_LMP`: the preconditioner is updated using Ritz LMP (*Gratton et al. 2011*).
- `Randomized_LMP`: the randomized LMP for the B-PCG is constructed at each GN step.

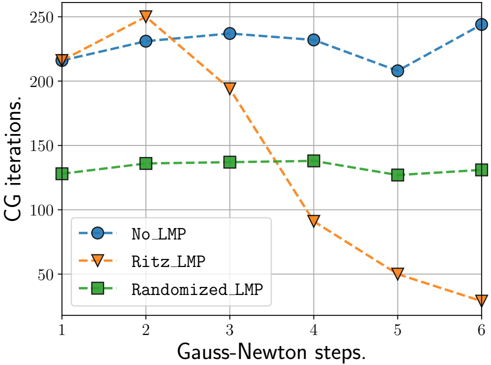
Results for HighObs





Results for HighObs

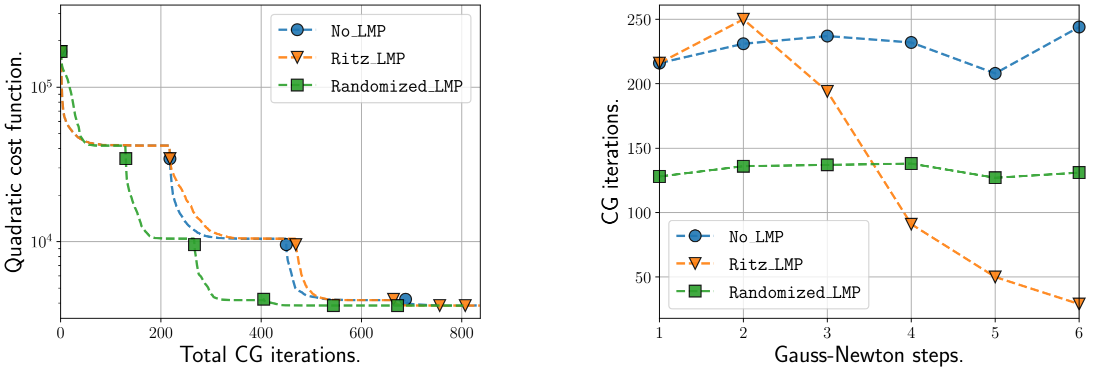




Observations:

- The gain obtained for `Randomized_LMP` seems fairly constant,
- The Ritz LMP is more efficient in the last GN steps due to the updates.

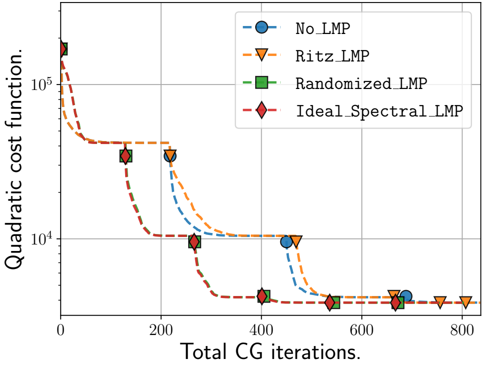
Results for HighObs

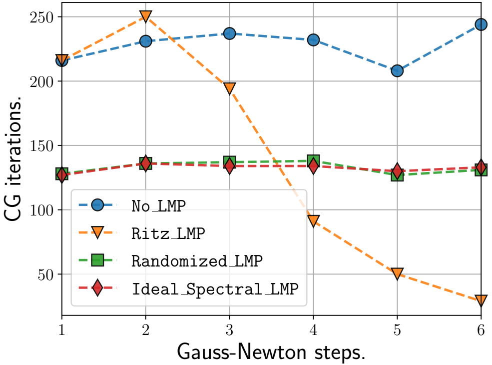


Observations:

- The gain obtained for Randomized_LMP seems fairly constant,
- The Ritz LMP is more efficient in the last GN steps due to the updates.

Results for HighObs

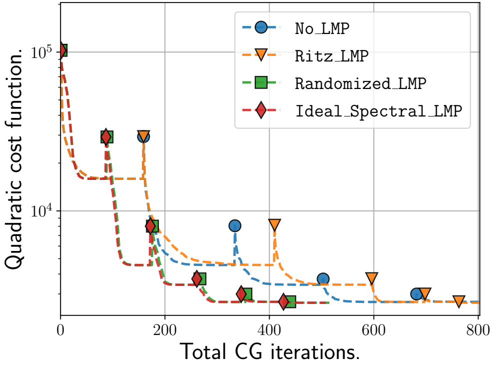


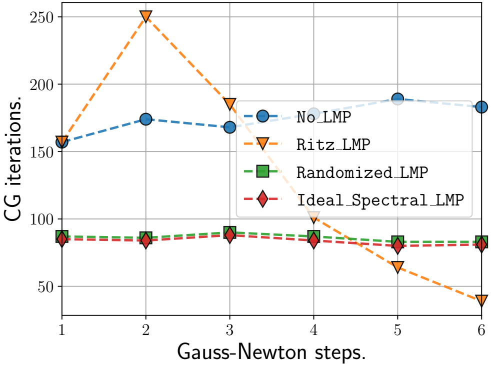


Observations:

- The gain obtained for `Randomized_LMP` seems fairly constant,
- The Ritz LMP is more efficient in the last GN steps due to the updates.

Results for MediumObs

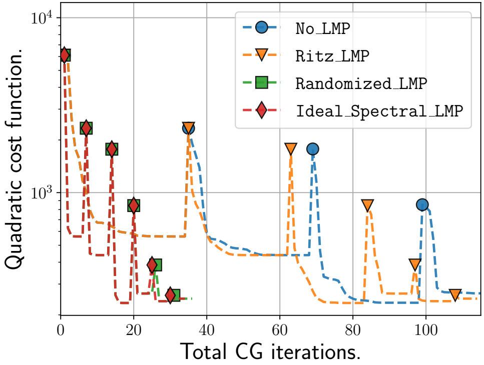


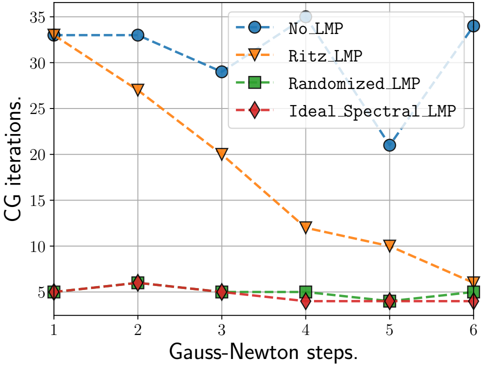


Observations:

- The gain obtained for `Randomized_LMP` seems fairly constant,
- The Ritz LMP is more efficient in the last GN steps due to the updates.

Results for LowObs





Observations:

- The gain obtained for Randomized_LMP seems fairly constant,
- The Ritz LMP is more efficient in the last GN steps due to the updates.

Overview of the computational costs for MediumObs

Accounting for the parallel nature of randomized methods, one has

	Ritz_LMP	Randomized_LMP
PCG iterations (total)	796	**524**
Storage (# vectors)	30 to 150	**50**

Additional construction cost	Ritz_LMP	Randomized_LMP
Applications of $R^{-1}, H_j, H_j^\mathsf{T}$	**0**	6
Applications of B	**0**	12

Outline

Conclusions and perspectives

Conclusions:

- We have proposed algorithms that generalize prior algorithms while improving the computational cost,
- We have derived an average-case analysis that is either new or improves state-of-the-art results,
- The numerical experiments conducted on a toy problem illustrated the behavior of the resulting preconditioners.

Perspectives:

- Study adaptive preconditioning strategies to combine randomized and Ritz approximations.
- Investigate the performance in larger scale applications (OOPS code from ECMWF).

References I

Youssef Diouane, Selime Gürol, Alexandre Scotto Di Perrotolo, and Xavier Vasseur.
A general error analysis for randomized low-rank approximation methods.
2022.

S. Gratton, A. S. Lawless, and N. K. Nichols.
Approximate Gauss–Newton Methods for Nonlinear Least Squares Problems.
SIAM Journal on Optimization, 18(1):106–132, January 2007.

S. Gratton, A. Sartenaer, and J. Tshimanga.
On A Class of Limited Memory Preconditioners For Large Scale Linear Systems With Multiple Right-Hand Sides.
SIAM Journal on Optimization, 21(3):912–935, July 2011.

Nathan Halko, Per-Gunnar Martinsson, and Joel A. Tropp.
Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions.
SIAM Review, 53(2):217–288, January 2011.

Jorge Nocedal and Stephen J. Wright.
Numerical Optimization.
Springer Series in Operation Research and Financial Engineering. Springer, New York, NY, second edition, 2006.

Arvind K. Saibaba, Jonghyun Lee, and Peter K. Kitanidis.
Randomized algorithms for generalized Hermitian eigenvalue problems with application to computing Karhunen–Loève expansion.
Numerical Linear Algebra with Applications, 23(2):314–339, March 2016.

Pseudocode of the algorithm

Input: B-symmetric matrices $\bar{A}, \bar{P} \in \mathbb{R}^{n\times n}$, integers $p, q \geq 1$ and $k \leq p$

% STEP 1

Draw a random matrix $\Omega \in \mathbb{R}^{n\times p}$, and set $V = \Omega$

Perform the QR factorization of $\bar{A}V = QR$ and set $X = Q$

for $j = 1, \ \ldots, \ q-1$ **do**

  Compute $V = \bar{P}X$

  Perform the QR factorization of $\bar{A}V = QR$ and set $X = Q$

end

% STEP 2

Form $T = R^{-\top}V^{\top}BX \in \mathbb{R}^{p\times p}$ and $\Phi = X^{\top}B\bar{P}X \in \mathbb{R}^{p\times p}$,

Solve the generalized Hermitian eigenvalue problem $TW = \Phi W\Theta$

Truncate W and Θ to keep k approximate eigenpairs.

Output: Matrices $\widetilde{V} = VW \in \mathbb{R}^{n\times k}$ and $\widetilde{\Lambda} = \Theta^{-1} \in \mathbb{R}^{k\times k}$ such that $\bar{P}\bar{A}\widetilde{V} \approx \widetilde{V}\widetilde{\Lambda}$.

Institut Supérieur de l'Aéronautique et de l'Espace

10 avenue Édouard Belin – BP 54032

31055 Toulouse Cedex 4 – France

Phone: +33 5 61 33 80 80

www.isae-supaero.fr