State-dependent Preconditioning for Data Assimilation
Victor Trappler, Arthur Vidard

To cite this version:
Victor Trappler, Arthur Vidard. State-dependent Preconditioning for Data Assimilation: Application to a Shallow Water Assimilation system. ISDA 2023 - 9th International Symposium on Data Assimilation, Oct 2023, Bologne, Italy. , pp.1-1, 2023. hal-04309242

HAL Id: hal-04309242
https://hal.science/hal-04309242
Submitted on 27 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
State-dependent Preconditioning for Data Assimilation

Victor Trappler 1,2  Arthur Vidard 1
1 AIRSEA Research team – INRIA – Laboratoire Jean Kuntzmann
2 Algèsim R&D Team – Eviène (ex Bull/Atos)

In short

Variational Data Assimilation requires the optimization of a high-dimensional function. This is usually done by linearizing the problem, and then inverting the Gauss-Newton matrix

\[ \mathbf{A}_x \]

Can we build a state-dependent preconditioner which would help for the inversion of the linear system?

\[ \mathbf{H} \]

ML framework

Objective

- Construct a preconditioner using DNN, which requires no call to \( \mathbf{A}_x \) when in use
- No access to \( \mathbf{A}_x \) during the training, no explicit construction of \( \mathbf{A}_x \)

Low-rank approximation

Let \( \mathbf{U}_o \in \mathbb{R}^{nxr} \) whose columns are orthonormal vectors, and \( \mathbf{A}_o = \text{diag}(\lambda_1, \ldots, \lambda_r) \in \mathbb{R}^{nxn} \)

\[ \mathbf{A}_o(x) = \mathbf{U}_o \mathbf{D}_o \mathbf{U}_o^T \]

Low-rank approximation is optimal wrt to the Frobenius norm

\[ \min_{\mathbf{F}} \| \mathbf{A}_x - \mathbf{A}_o(x) \|_F^2 \]

but \( \mathbf{A}_x \) cannot be explicitly constructed in practice: estimation of F-norm

\[ \mathbf{L}(\theta, x) = \frac{1}{n} \sum_{i=1}^{n} \| \mathbf{A}_x \mathbf{z}_i - \mathbf{A}_o(x) \mathbf{z}_i \|_2^2 \quad \text{with} \quad \mathbf{z}_i \sim \mathcal{N}(0, I_n) \]

Possibility of online training:

- Only \( \mathbf{A}_x \) needed instead of constructing and storing \( \mathbf{A}_x \) explicitly

When trained, we can construct the preconditioner as (15)

\[ \mathbf{L}_o = \mathbf{U}_o \mathbf{D}_o \mathbf{U}_o^T + \mathbf{L}, \quad \text{sp} \mathbf{L}_o = \{1, \ldots, 1, \lambda_1^{-1}, \ldots, \lambda_r^{-1} \} \]

Numerical Results

SW assimilation system, \( n = 120000 \), UNet-like architecture: \( \varrho, \nu, v \) as features

State Dependent Preconditioner

Preconditioning

Instead of solving \( \mathbf{A}_x \mathbf{dx} = \mathbf{b}_o \), solve \( (\mathbf{L}^T \mathbf{L}) \mathbf{dx} = \mathbf{L}^T \mathbf{b} \) and \( \mathbf{dx} = \mathbf{L} \mathbf{z} \)

- \( \mathbf{H} = \mathbf{L} \mathbf{L}^T \) symmetric, positive definite, cheap to compute and to apply
- \( \mathbf{H} \) should be close to \( \mathbf{A}_x^{-1} \)
- \( 1 \leq \kappa(\mathbf{A}_x) \leq \kappa(\mathbf{A}_x) \)

"One-fits-all" preconditioner do not exist, most include information on spectrum of \( \mathbf{A}_x \), which depends on \( x \)

State-dependent preconditioner

We propose to construct a mapping

\[ x \rightarrow \mathbf{H}(x) \]

where \( \mathbf{H}(x) \) is a preconditioner well-suited for the linear system \( \mathbf{A}_x \mathbf{dx} = \mathbf{b}_o \)

Challenges

- \( \mathbf{H}(x) \in \mathbb{R}^{nxn} \) is spd (ie \( n(n+1)/2 \) "free" parameters)
- \( \mathbf{A}_x \) is not stored explicitly (only accessible as \( \mathbf{L}^T \mathbf{L} \), \( \mathbf{A}_x \) and high-dimensional
- Independence with respect to the observations (thus to \( \mathbf{b}_o \))
- \( \mathbf{H}(x) \) should contain spectral information of \( \mathbf{A}_x \)

References


