Adaptive Preconditioning for TV Regularisation

Malena Sabaté Landman
Silvia Gazzola

Como, 23rd May 2018

Smoothing Norm Preconditioning
[P. C. Hansen and T. K. Jensen (2007)]

Tikhonov Regularization

$$\min_{x \in \mathbb{R}^n} \{||Ax - b||_2^2 + \lambda ||Lx||_2^2\} \quad \text{for} \quad L \in \mathbb{R}^{n \times n}$$

Standard form transformation

$$\min_{y \in \mathbb{R}^n} \{||\bar{A}y - \bar{b}||_2^2 + \lambda ||y||_2^2\}$$

for $\hat{A} = AL^{-1}$, $\hat{b} = b$ and $y = Lx$.

Smoothing Norm Preconditioning

Tikhonov Regularization

$$\min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda||Lx||_2^2\} \quad \text{for} \quad L\in\mathbb{R}^{p\times n}$$

Standard form transformation

$$\min_{y\in\mathbb{R}^n}\{||\bar{A}y-\bar{b}||_2^2+\lambda||y||_2^2\}$$

for $\hat{A}=AL_A^{\dagger}$, $\hat{b}=b-Ax_0$ and $x_L=L_A^{\dagger}y_L+x_0=L_A^{\dagger}y_L+Kt_0$ for $\mathcal{R}(K)=\mathcal{N}(L)$. We can prove that $x_0=K(AK)^{\dagger}b$.

In order to apply GMRES we modify the system as:

$$Ax_L = b \rightarrow A\left[L_A^\dagger,\ K\right]\left[\begin{array}{c}\bar{y}_L\\ t_0\end{array}\right] = b\,,$$

once premultiplied by $[L_A^\dagger,\ K]^T$, gives the 2×2 block system

$$\left[\begin{array}{cc}(L_A^\dagger)^T A L_A^\dagger & (L_A^\dagger)^T AK\\ K^T A L_A^\dagger & K^T AK\end{array}\right]\left[\begin{array}{c}\bar{y}_L\\ t_0\end{array}\right] = \left[\begin{array}{c}(L_A^\dagger)^T b\\ K^T b\end{array}\right].$$

We can easily eliminate t_0 from this system by inverting the 1×1 $(2,2)$ block, obtaining the Schur complement system,

$$(D^\dagger)^T PAL_A^\dagger \bar{y}_L = (D^\dagger)^T Pb \rightarrow (D^\dagger)^T PA\bar{x}_L = (D^\dagger)^T Pb\,,$$

where $P = I - AK(K^T AK)^{-1}K^T \in \mathbb{R}^{N\times N}$.

TV-ReGMRES(WD)

Total variation Regularization

$$\min_{x \in \mathbb{R}^n} \{||Ax - b||_2^2 + \lambda TV(x)\}$$

Discretization of TV regularization term

- 1d case. In a discrete setting with $x \in \mathbb{R}^N$, $\mathsf{TV}(x) = \|D_{1d}x\|_1$,

$$D_{1d} = \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(N-1)\times N}.$$

- 2d case. In a discrete setting,

$$\mathsf{TV}(x) = \| \left((D^h x)^2 + (D^v x)^2 \right)^{1/2} \|_1 \,,$$

where, $x \in \mathbb{R}^N$ for an array $V \in \mathbb{R}^{n\times n}$ so $N = n^2$ and the discrete first derivatives in the horizontal and vertical directions are given by

$$D^h = (D_{1d} \otimes I) \in \mathbb{R}^{n(n-1)\times n^2} \;,\; D^v = (I \otimes D_{1d}) \in \mathbb{R}^{n(n-1)\times n^2}$$

Iterative Reweighed Norm
[B. Wohlberg and P. Rodríguez (2007)]

$$||x||_1 \approx ||Wx||_2$$

- 1d case. The weighting matrix is:

$$W_{1d} = W_{1d}(x) = \text{diag}\left(|D_{1d}x|^{-1/2}\right) \in \mathbb{R}^{(N-1)\times(N-1)},$$

  - Component-wise modulus and exponentiation.

- 2d case. Less straightforwardly:

$$\begin{aligned}
\tilde{W}_{2d} &= \tilde{W}_{2d}(x) = \text{diag}\left(\left((D^h x)^2 + (D^v x)^2\right)^{-1/4}\right) \in \mathbb{R}^{\tilde{N}\times\tilde{N}}, \\
W_{2d} &= W_{2d}(x) = \begin{bmatrix} \tilde{W}_{2d} & 0 \\ 0 & \tilde{W}_{2d} \end{bmatrix} \in \mathbb{R}^{2\tilde{N}\times 2\tilde{N}}, \\
D_{2d} &= \begin{bmatrix} D^h \\ D^v \end{bmatrix} \in \mathbb{R}^{2\tilde{N}\times N} \text{ where } \tilde{N} = n(n-1).
\end{aligned}$$

TV-ReGMRES(WD)

Total variation Regularization

$$\min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda TV(x)\} = \min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda||Dx||_1\}$$
$$\approx \min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda||WDx||_2\}$$

TV-ReGMRES(WD)

Total variation Regularization

$$\min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda TV(x)\} = \min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda||Dx||_1\}$$
$$\approx \min_{x\in\mathbb{R}^n}\{||Ax-b||_2^2+\lambda||WDx||_2\}$$

Standard form transformation

$$\min_{y\in\mathbb{R}^n}\{||\bar{A}y-\bar{b}||_2^2+\lambda||y||_2^2\}$$

for $\hat{A} = A(WD)_A^{\dagger}$, $\hat{b} = b - Ax_0$ and
$x_L = (WD)_A^{\dagger}y_L + x_0 = (WD)_A^{\dagger}y_L + Kt_0$
for $\mathcal{R}(K) = \mathcal{N}(WD)$ and $x_0 = K(AK)^{\dagger}b$.

We can reproduce the results we had for the SN-GMRES:

$$A\left[L_A^{\dagger},\ K\right]\left[\begin{array}{c} y_L \\ t_0 \end{array}\right] = b\,,$$

once premultiplied by $[D^{\dagger},\ K]^T$, gives the 2×2 block system

$$\left[\begin{array}{cc} (D^{\dagger})^T A L_A^{\dagger} & (D^{\dagger})^T A K \\ K^T A L_A^{\dagger} & K^T A K \end{array}\right]\left[\begin{array}{c} y_L \\ t_0 \end{array}\right] = \left[\begin{array}{c} (D^{\dagger})^T b \\ K^T b \end{array}\right].$$

We can easily eliminate t_0 from this system by inverting the 1×1 $(2,2)$ block, obtaining the Schur complement system,

$$(D^{\dagger})^T P A L_A^{\dagger} y_L = (D^{\dagger})^T P b \rightarrow (D^{\dagger})^T P A \bar{x}_L = (D^{\dagger})^T P b\,,$$

where $P = I - AK(K^T AK)^{-1} K^T \in \mathbb{R}^{N \times N}$.

The TV-ReGMRES(v) method.

Input: $\widehat{A}$, $\widehat{b}$, $\{i_1, ..., i_k\}$

1. Perform $m^{(1)}$ GMRES(D) iterations with input $\widehat{A}$ and $\widehat{b}$ to obtain the relation $\widehat{A}Z_m = V_{m+1}G_m$
2. Find $\bar{x}_L = Z_m s_m$ and $r = V_{m+1}(||\widehat{b}||_2 e_1 - G_m s_m)$, where $s_m = \arg\min_{s\in\mathbb{R}^m} \|G_m s - ||\widehat{b}||_2 e_1\|_2$
3. Compute the reduced QR factorization $G_m P_k = Q_{\mathsf{GP}} R_{\mathsf{GP}}$
4. Set $C_k = V_{m+1} Q_{\mathsf{GP}}$ and $U_k = Z_m P_k (R)^{-1}_{\mathsf{GP}}$

For $j = 1, 2, \ldots$ until a stopping criterion is satisfied

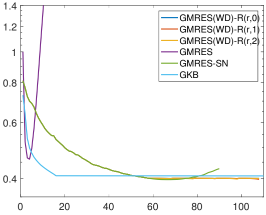
5. Perform m GMRES(WD) iterations with input $(I - C_k(C_k)^T)\widehat{A}$ and $(I - C_k(C_k)^T)v$ use recycling strategies to compute an expression $\widehat{A}\tilde{Z}_{m+k} = \tilde{V}_{m+k+1}\tilde{G}_{m+k}$ and compute an approx. of $\bar{x}_L$
6. Compute the reduced QR factorization $G_{k+m} P_k = Q_{\mathsf{GP}} R_{\mathsf{GP}}$.
7. Set $C_k = \tilde{V}_{m+1} Q_{\mathsf{GP}}$ and $U_k = \tilde{Z}_{k+m} P_k (R_{HP})^{-1}$

end

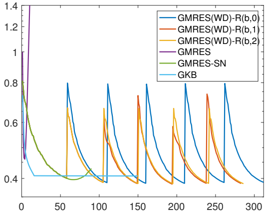
Take $x_L = \bar{x}_L + x_0$

Numerical Experiments

History of relative errors

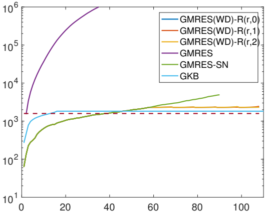


Restarting r

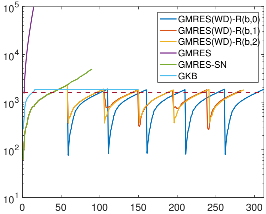


Restarting b

History of relative total variation



Restarting r

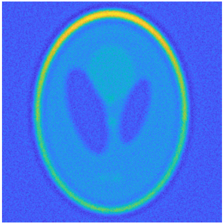


Restarting b

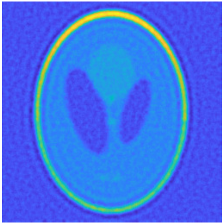
Numerical Experiments

Best reconstructions for different methods

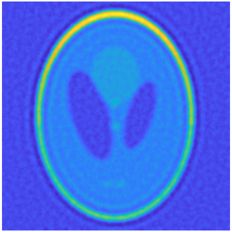
GKB



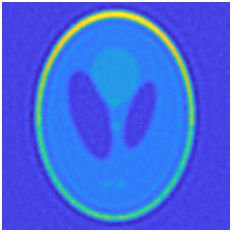
GMRES



TV-GMRES(D)



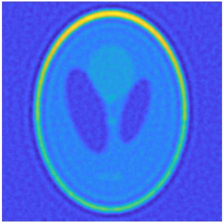
TV-ReGMRES(WD)-R(r)



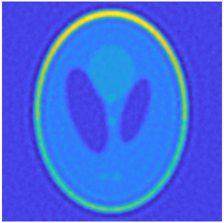
TV-ReGMRES(WD)-R(b)

Numerical Experiments

Best reconstructions for different outer iterations of TV-ReGMRES(WD)-R(b)



Iteration 1



Iteration 2

Iteration 3

Selected references

B. Wohlberg and P. Rodríguez (2007). An iteratively reweighted norm algorithm for minimization of total variation functionals. *IEEE Signal Process. Lett.*, 14:948–951.

P. C. Hansen and T. K. Jensen (2007). Smoothing-norm preconditioning for regularizing minimum-residual methods. *SIAM Journal on Matrix Analysis and Applications*, 29(1):1–14.