Krylov-Simplex method for inverse problems

Wim Vanroose Jeffrey Cornelis

Dept of Mathematics, Universiteit Antwerpen, Belgium

Minimisation in Different Norms

Least Squares

$$\min \|b - Ax\|_2 = \min \sqrt{\sum_{i=1}^{m} (b - Ax)_i^2} \quad (1)$$

Least absolute values

$$\min \|b - Ax\|_1 = \min \sum_{i=1}^{m} |(b - Ax)_i| \quad (2)$$

least absolute deviation

$$\min \|b - Ax\|_\infty = \min \max_{i=1\ldots m} |(b - Ax)_i| \quad (3)$$

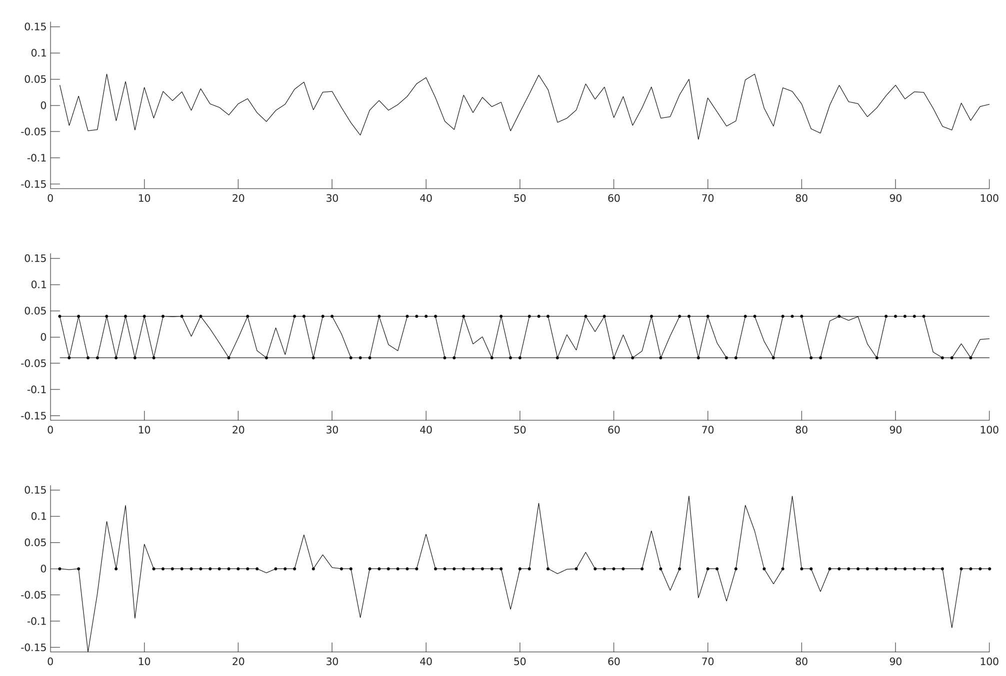

Figure: We compare the residuals after optimization of x in the $\|\cdot\|_2$, $\|\cdot\|_1$ and $\|\cdot\|_\infty$

Krylov subspace Methods

We start with an initial guess x_0 and calculate the initial residual $r = b - Ax_0$. A Krylov subspace is then

$$\mathcal{K}_k(A, r_0) = \text{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0\}, \quad (4)$$

Conjugate Gradients

$$\min_{x \in x_0 + \mathcal{K}_k(A, r_0)} \|x - x^*\|_A \quad \rightarrow \quad T_{kk} y_k = \|r_0\| e_1 \quad (5)$$

where $x = x_0 + V_k y_k$ with V_k a basis for the Krylov subspace

Generalized Minimal Residual (GMRES)

$$\min_{x \in x_0 + \mathcal{K}_k(A, r_0)} \|b - Ax\|_2 \quad \rightarrow \quad \min \| \|r\|_2 e_1 - H_{k+1,k} y_k \| \quad (6)$$

Krylov-Simplex for ℓ_∞

Definition (max-norm Krylov)

Let $A \in \mathbb{R}^{m \times n}$ a matrix and $b \in \mathbb{R}^m$ a right hand side. The iterates of the ℓ_∞-norm Krylov-Simplex are given by

$$x_k := \text{argmin}_{x \in x_0 + \mathcal{K}_k} \|r_k\|_\infty. \quad (7)$$

This corresponds to the optimization problem

$$\begin{aligned} &\min \gamma_k \\ &\text{s.t.} \ -\gamma_k \mathbb{1}_m \le r_0 - AV_k y_k \le \gamma_k \mathbb{1}_m, \end{aligned} \quad (8)$$

The KKT conditions

$$\begin{aligned} \mathbb{1}_m^T(\lambda + \mu) &= 1, && (9)\\ V_k^T A^T(\lambda - \mu) &= 0, && (10)\\ -\gamma_k \mathbb{1}_m \le r_0 - AV_k y_k &\le \gamma_k \mathbb{1}_m && (11)\\ \lambda \ge 0, \quad \mu &\ge 0, && (12)\\ \lambda_i(r_0 - AV_k y + \gamma_k \mathbb{1}_m)_i &= 0, \quad \forall i \in \{1..m\} && (13)\\ \mu_i(\gamma_k \mathbb{1}_m - (r_0 - AV_k y))_i &= 0, \quad \forall i \in \{1..m\}. && (14) \end{aligned}$$

The first two equations, (9) and (10) are the dual equations, the third equation is the primal equation and the last two, (13) and (14), are the complementarity conditions that couple all equations together.

Definition (Basic set)

A basic set $\mathcal{B}_k = \mathcal{B}_{k,<} \cup \mathcal{B}_{k,>}$ for $\mathcal{K}_k$ and $\gamma_k > 0$ is a subset of $k+1$ indices from $\{1, \ldots, m\}$ such that

$$(r_0 - AV_k y_k)_i = -\gamma_k \quad \forall i \in \mathcal{B}_{k,<}, \quad (15)$$

and

$$(r_0 - AV_k y_k)_i = \gamma_k \quad \forall i \in \mathcal{B}_{k,>}. \quad (16)$$

and for all $i \in \mathcal{N}_k := \{1, \ldots, m\} \setminus \mathcal{B}_k$ holds that

$$|(r_0 - AV_k y_k)_i| \le \gamma_k \quad \forall i \in \mathcal{N}_k \quad (17)$$

and

$$B_k := \begin{pmatrix} -\mathbb{1}_{\mathcal{B}_{k,<}} & AV_k|_{\mathcal{B}_{k,<}} \\ \mathbb{1}_{\mathcal{B}_{k,<}} & AV_k|_{\mathcal{B}_{k,>}} \end{pmatrix} \quad (18)$$

with $B_k \in \mathbb{R}^{k+1 \times k+1}$ non-singular.

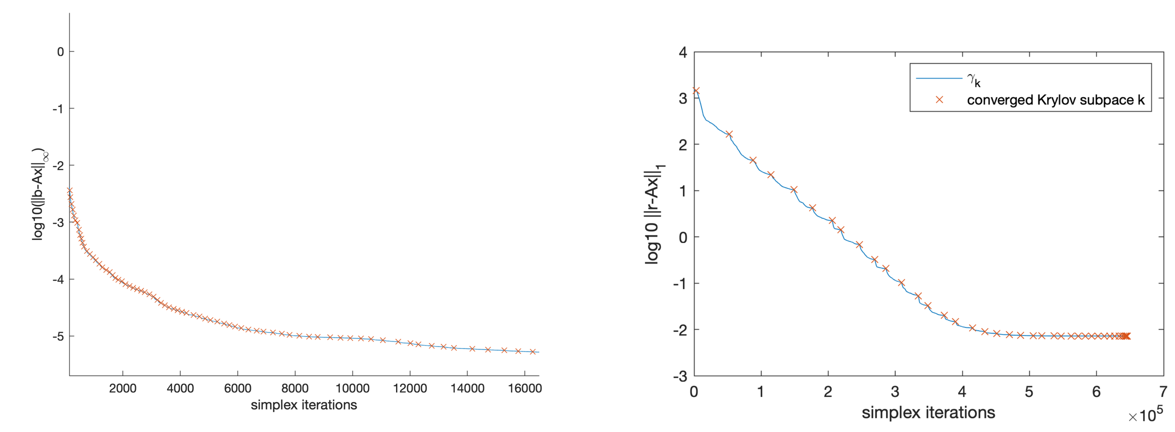

Figure: Left: The $\|b - Ax\|_\infty$ as a function of the simplex iterations. Each cross indicates an optimal point within the Krylov subspace. Right: the $\|b - Ax\|_1$ as a function of the simplex iterations.

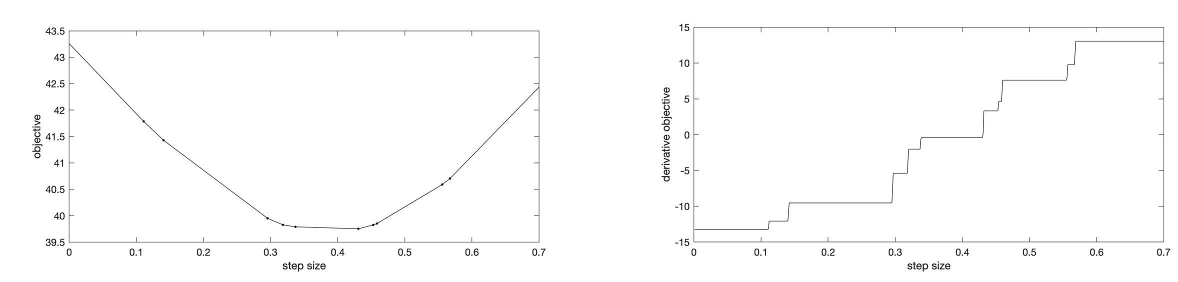

Figure: We show, left, an objective involving an ℓ_1-norm along a search direction p for various step sizes s. We have a piecewise linear function with breakpoints. The slope changes occur when one particular component of the vector in the ℓ_1-norm switches sign, as the step size changes. The derivative, right, is a piecewise constant function.

Krylov-Simplex for ℓ_1

Definition (ℓ_1-norm Krylov)

Let $A \in \mathbb{R}^{m \times n}$ a matrix and $b \in \mathbb{R}^m$ a right hand side. The iterates of the ℓ_1-norm Krylov-Simplex are given by

$$x_k := \text{argmin}_{x \in x_0 + \mathcal{K}_k} \|r_k\|_1. \quad (19)$$

This corresponds to the optimization problem

$$\begin{aligned} &\min \mathbb{1}_m^T \gamma_k, \\ &\text{s.t} \ -\gamma_k \le r_0 - AV_k y_k \le \gamma_k. \end{aligned} \quad (20)$$

KKT conditions

$$\begin{aligned} \lambda + \mu &= 1, && (21)\\ V_k^T A^T(\lambda - \mu) &= 0, && (22)\\ -\gamma_k \le (r_0 - AV_k y_k) &\le \gamma_k, && (23)\\ \mu \ge 0, \quad \lambda &\ge 0, && (24)\\ \lambda_i((r_0 - AV_k y_k) + \gamma_k)_i &= 0, \quad \forall i \in \{1..m\} && (25)\\ \mu_i(\gamma_k - (r_0 - AV_k y_k))_i &= 0, \quad \forall i \in \{1..m\} && (26) \end{aligned}$$

The first two equations, (21) and (22), are the dual equations. The third is the primal equation. The last two, (25) and (26), are the complementarity conditions. Note the similarity with the KKT conditions (9)-(14).

Rank-1 updates in Simplex

When the basic set is updated we have the following update for

$$B_k := ((\pm 1)|_{\mathcal{B}_k} \ AV_k|_{\mathcal{B}_k}) + e_j \left[((\pm 1)_r \ (AV_k)_r) - ((\pm 1)_q \right. \quad (27)$$

where j is the position of q in the set $\mathcal{B}_j$. This can be implemented with `QRupdate`.

Convergence properties

Since

$$\begin{aligned} &\|b - Ax_k^{\text{KS}_\infty}\|_\infty \le \|b - Ax_k^{\text{GMRES}}\|_\infty \\ &\le \|b - Ax_k^{\text{GMRES}}\|_2 \le \|b - Ax_k^{\text{KS}_\infty}\|_2 \end{aligned} \quad (28)$$

We conclude that

$$\|b - Ax_k^{\text{KS}_\infty}\|_\infty \le \|b - Ax_k^{\text{GMRES}}\|_2 \quad (29)$$

So a good preconditioner for GMRES is also a good preconditioner for Krylov-Simplex.

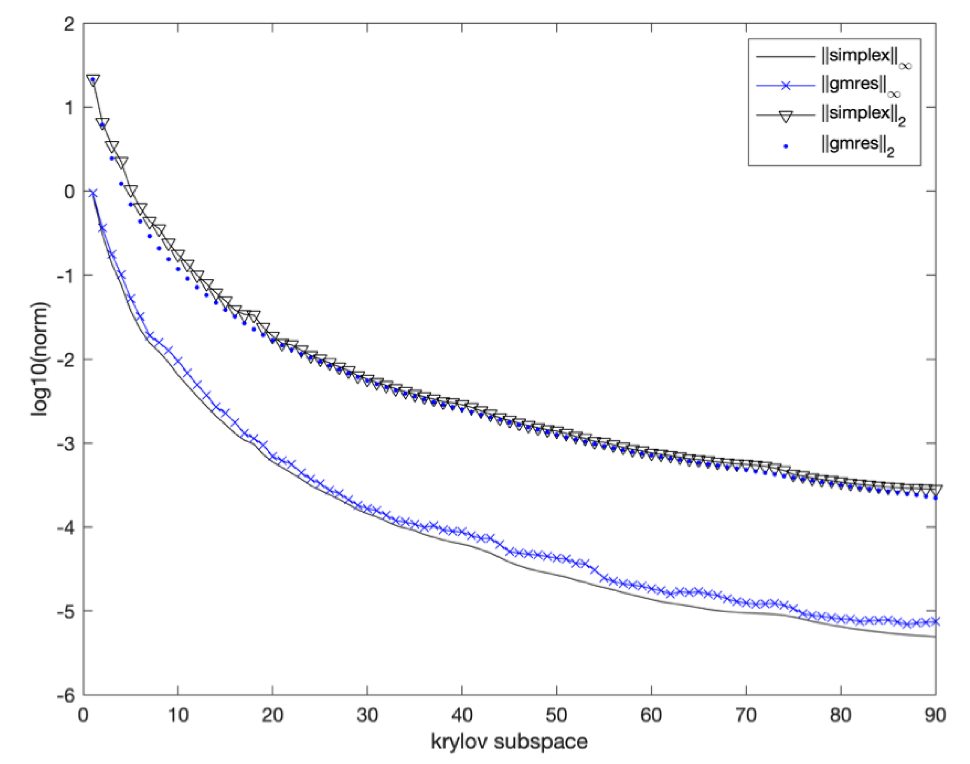

Figure: Comparison of the convergence of Krylov-Simplex versus GMRES over the same Krylov subspace.

Conclusions and future steps

- Algorithm can be written with only rank-one updates.
- Expensive since the solutions are constructed each iteration.

Future:

- Comparision with column generation.
- Long-step Simplex with bound flipping.
- Extension $\ell_p - \ell_q$ inverse problems

References

Vanroose, Wim, and Jeffrey Cornelis. ”Krylov-Simplex method that minimizes the residual in ℓ_1-norm or ℓ_∞-norm.” arXiv preprint arXiv:2101.11416 (2021).

Algorithm 3.1 Krylov-Simplex for $\min_{x \in x_0 + \mathcal{K}_k} \|b - Ax\|_\infty$

$r_0 = b - Ax_0$
$\gamma_0, i = \max_i |(r_0)_i|$
$\mathcal{B}_0 = \{i\}$, index where the max is reached.
$V_1 = [r_0/\|r_0\|]$
for $k = 1, \ldots,$maxit **do**
  Calculate $AV_k = [AV_{k-1} Av_k]$ and store and update $V_{k+1} = [V_k, v_{k+1}]$.
  if $k = 1$ **then**
    $B_1 = (\pm 1)$,
    $[Q, R] = \text{qr}(B_1)$.
  end if
  Solve $B_k d = -Av_{k+1}|_{\mathcal{B}_k}$ using Q and R.
  Set $d := [d(2:end)/d1; 1/d(1)]$
  Set $y_k = [y_k; 0]$
  $r_k, \Delta\gamma, y_k, r$ = StepSizeKrylovExpansion ().
  $\mathcal{B}_k = \mathcal{B}_{k-1} \cup \{r\}$,
  while $l = 1, \ldots,$ maxsimplexiters **do**
    if $l = 1$ **then**
      Update the Q and R using (3.45).
    else
      Update the Q and R using (3.44).
    end if
    Solve $B_k^T \begin{pmatrix} \lambda_{\mathcal{B}_<} \\ -\mu_{\mathcal{B}_>} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ using Q and R.
    if $\lambda \ge 0$ and $\mu \ge 0$ **then**
      Break; Solution Found.
    else
      $q = \min(\lambda_i, \mu_i)$ leaving index,
      $\mathcal{B}_k = \mathcal{B}_k \setminus \{q\}$,
      Solve $B_k d = e_q$ using Q and R,
      Set $d := d(2:end)$, Set $d_1 = d(1)$,
      $r_k, \Delta\gamma, y_k, r$ = StepSizeSimplex (),
      $\mathcal{B}_k = \mathcal{B}_l \cup \{r\}$.
    end if
  end while
  $x_k = x_0 + V_k y_k$.
  $\|r_k\|_\infty = \gamma_k$.
end for