Enriched Krylov Subspace Methods
Ill-Posed Problems

Daniela Calvetti
Lothar Reichel
Abdallah Shuibi

http://lanczos.cwru.edu/ dxc57

Noise-Free Problem

Linear system

$$Ax = b \qquad A \in R^{m \times n} (1)$$

Noisy Problem

$$Ax = b^\delta \qquad (2)$$

where

$b^\delta = b + e$; e=noise; $||e|| = \delta$; δ known.

Consider the Normal Equations associated with (2)

$$A^T A x = A^T b^\delta.$$

Assume:

- n, m are very large;
- A is of ill-determined rank;
- δ is known.

Popular solution methods

1. Conjugate Gradient method applied to normal equations

$$A^T A x = A^T b^\delta$$

2. Tikhonov regularization:

$$\min_{x \in R^n} (\|Ax - b^\delta\|^2 + \mu \|x\|^2)$$

Normal equations for Tikhonov regularization:

$$(A^T A + \mu I)x = A^T b^\delta$$

can be solved by CG method without forming $A^T A$.

Recall:

kth iterate x_k of CGNR method ($x_0 = 0$)

belongs to Krylov subspace

$$\mathcal{K}_k(A^TA, A^Tb^\delta) := span\{A^Tb^\delta, (A^TA)A^Tb^\delta, \dots, (A^TA)^{k-1}A^Tb^\delta\} \tag{3}$$

and is characterized by

$$\Phi(x_k) = \min_{x \in \mathcal{K}_k(A^TA, A^Tb^\delta)} \Phi(x)$$

where

$\Phi(x) := \frac{1}{2}x^TA^TAx - x^TA^Tb^\delta$.

Determination of the minimizer x_k

Sequential linear searches along
$A^T A$-conjugate directions $p_0, p_1, \ldots, p_{k-1}$
that span $\mathcal{K}_k(A^T A, A^T b^\delta)$.

Determine x_k from x_{k-1} and p_{k-1}
according to

$$x_k := x_{k-1} + \alpha_{k-1} p_{k-1}$$

where $\alpha_{k-1} \in R$ solves

$$\min_{\alpha \in R} \Phi(x_{k-1} + \alpha p_{k-1}).$$

Introduce the residual error associated with x_k:

$$r_k := A^T b^\delta - A^T A x_k.$$

Then

$$p_k := r_k + \beta_{k-1} p_{k-1}$$

Choose β_{k-1} so that p_k is $A^T A$ conjugate to the previous search directions.

Discrepancy

$$d_k := b^\delta - Ax_k$$

.

Then $r_k = A^T d_k$.

It was shown by Hestenes and Stiefel that

$$\|d_{k+1}\| \leq \|d_k\|; \qquad \|x_{k+1}\| \geq \|x_k\|$$

Discrepancy principle:

Terminate CGNR iterates as soon as

$$\|d_k\| \leq \delta.$$

Algorithm (CGNR)

$x_0 := 0; d_0 = b^\delta; r_0 = A^* b^\delta;$

$p_0 = r_0; t_0 = A p_0;$

for $k = 1, 2, \ldots$ until stopping criterion is satisfied

$$\alpha_k = \|r_{k-1}\|^2 / \|t_{k-1}\|^2$$
$$x_k = x_{k-1} + \alpha_k p_{k-1};$$
$$d_k = d_{k-1} - \alpha_k t_{k-1};$$
$$r_k = A^* d_k;$$
$$\beta_k = \|r_k\|^2 / \|r_{k-1}\|^2;$$
$$p_k = r_k + \beta_k p_{k-1};$$
$$t_k = A p_k$$

Enriched CGNR method Let

$$\mathcal{Q} := span\{q^{(1)}, \ldots, q^{(\ell)}\}$$

be an ℓ-dimensional vector space.

The **enriched CGNR method** determines the kth iterate $\tilde{x}_k$ in

$$\mathcal{K}_k(A^T A, A^T b^\delta) \cup \mathcal{Q}$$

Note: the improvement over CGNR is determined by how "pertinent" the subspace $\mathcal{Q}$ is and how much its information is already in $\mathcal{K}_k(A^T A, A^T b^\delta)$.

Let

$$\{\tilde{q}_k^{(1)}, \tilde{q}_k^{(2)}, \ldots, \tilde{q}_k^{(\ell)}\} = span\,\mathcal{Q}\backslash\mathcal{K}_k(A^TA, A^Tb^\delta)$$

and assume $\tilde{q}_k^{(j)}$ are A^TA-conjugate to search directions $p_0, p_1, \ldots, p_{k-1}$.

Set $\tilde{q}_0^{(j)} = q^{(j)}$.
For $k > 0$ compute

$$\tilde{q}_k^{(j)} := \tilde{q}_{k-1}^{(j)} - \gamma_{k-1}^{(j)} p_{k-1}, \qquad 1 \leq j \leq \ell.$$

Choose $\gamma_{k-1}^{(j)}$ so that $\tilde{q}_k^{(j)}$ are A^TA-conjugate to p_{k-1}, for $1 \leq j \leq \ell$.

Introduce the matrix

$$\tilde{Q}_k = [\tilde{q}_k^{(1)}, \tilde{q}_k^{(2)}, \dots, \tilde{q}_k^{(\ell)}] \in R^{n\times \ell}$$

The kth iterate $\tilde{x}_k$ determined by enriched CGNR method is

$$\tilde{x}_k := x_k + \tilde{Q}_k \tilde{y}_k = x_{k-1} + \alpha_{k-1} p_{k-1} + \tilde{Q}_k \tilde{y}_k$$

where $\tilde{y}_k$ solves minimization problem

$$\min_{\tilde{y}\in R^\ell} \Phi(x_k + \tilde{Q}_k \tilde{y}).$$

Compute $\tilde{y}_k$ as solution of least squares problem

$$\min_{\tilde{y}\in R^n} \|A\tilde{Q}_k \tilde{y} - d_k\|$$

Theorem 1
The discrepancies $\tilde{d}_k := b^\delta - A\tilde{x}_k$
associated with the iterates $\tilde{x}_k$ satisfy

$$\|\tilde{d}_k\| \leq \|\tilde{d}_{k-1}\| \qquad k = 1, 2, \ldots$$

Note: The norm of the iterates $\tilde{x}_k$ is not
an increasing function of k.

Algorithm: (Enriched CGNR algorithm)

$d_0 := b^\delta - Ax_0;\;\; \tilde{d}_0 := d_0;$

$r_0 := A^T d_0;\;\; p_0 := r_0;\;\; V_0 = A\tilde{Q}_0;\;\; S_0 = A^T V_0$

for $k = 0, 1, 2, \ldots,$ **until** stopping criterion satisfied

1. $w_k := Ap_k$
2. $\alpha_k := \frac{r_k^T r_k}{w_k^T w_k}$
3. $x_{k+1} := x_k + \alpha_k p_k$
4. $d_{k+1} := d_k - \alpha_k w_k$
5. $f_k := A^T w_k$
6. $r_{k+1} := r_k - \alpha_k f_k$
7. $\beta_k := \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$
8. $p_{k+1} := r_{k+1} + \beta_k p_k$
9. $g_k := \frac{V_k^T w_k}{w_k^T w_k}$
10. $\tilde{Q}_{k+1} := \tilde{Q}_k - p_k g_k^T$
11. $V_{k+1} := V_k - w_k g_k^T$

12. Solve the least-squares problem
$\min_{\tilde{y} \in R^n} \|V_{k+1}\tilde{y} - d_{k+1}\|$ for $\tilde{y}_{k+1}$

13. $\tilde{x}_{k+1} := x_{k+1} + \tilde{Q}_{k+1}\tilde{y}_{k+1}$

14. $\tilde{d}_{k+1} := d_{k+1} - V_{k+1}\tilde{y}_{k+1}$

15. $S_{k+1} := S_k - f_k g_k^T$

16. $\tilde{r}_{k+1} := r_{k+1} - S_{k+1}\tilde{y}_{k+1}$

end k

Note: can solve least squares problem by QR factorization of V_{k+1}, obtainable by updating QR factorization of V_k.

Enriched CG method for penalized least-squares problems

1. Formulate the problem as

$$\min_{\tilde{y}\in R^{\ell}} \left\| \begin{bmatrix} V_{k+1} \\ \mu^{1/2}\tilde{Q}_{k+1} \end{bmatrix} \tilde{y} - \mu^{1/2} \begin{bmatrix} 0 \\ r_{k+1} \end{bmatrix} \right\|.$$

2. Use QR factorization of V_k and $\tilde{Q}_k$ to compute QR factorization of V_{k+1} and $\tilde{Q}_{k+1}$

3. Reduce the problem to a minimization of small size.

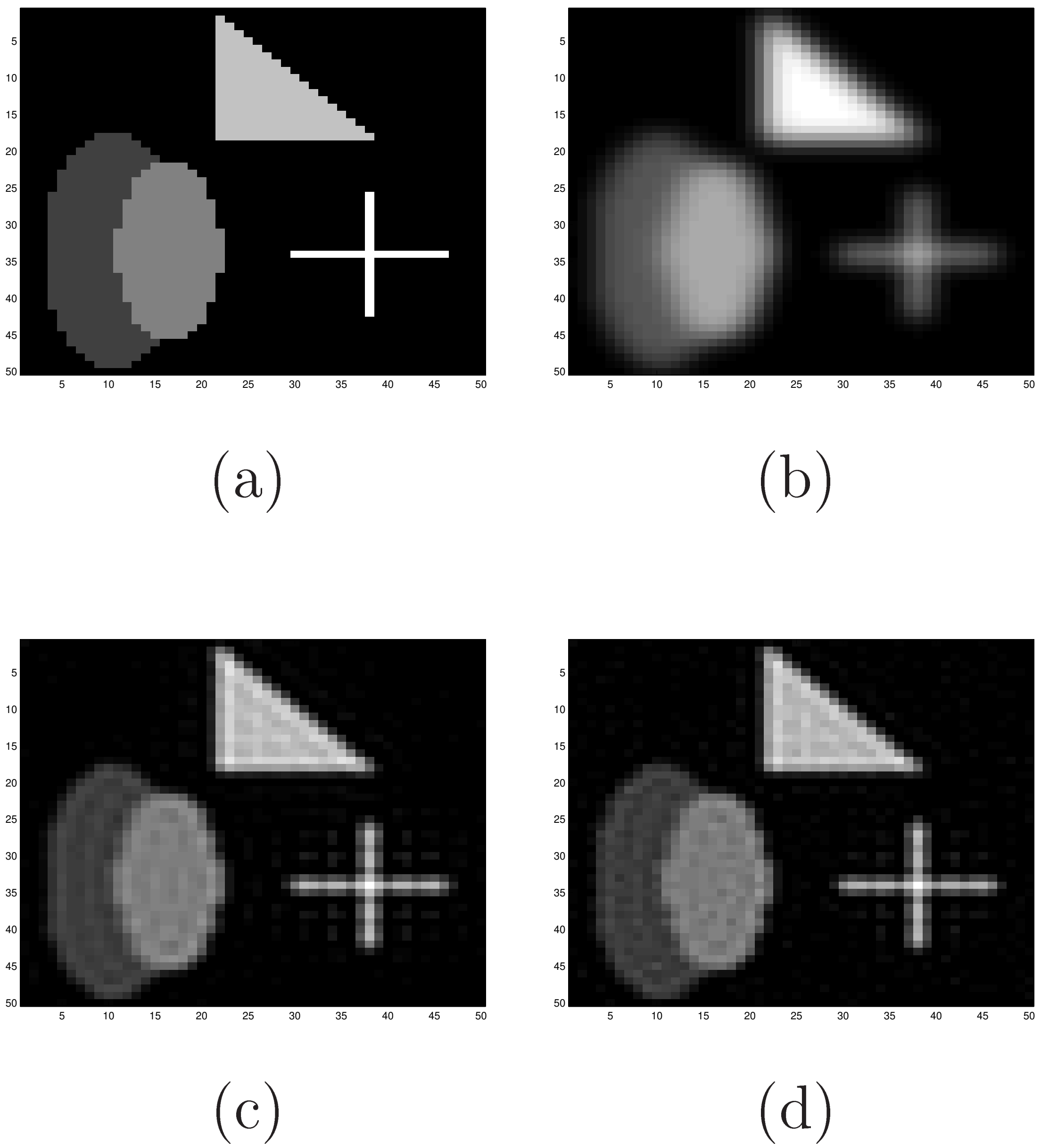
Algorithm 2: Enriched CG for Tikhonov

1. $w_k := Ap_k$
2. $\alpha_k := \frac{r_k^T r_k}{w_k^T w_k + \mu p_k^T p_k} \leftarrow$
3. $x_{k+1} := x_k + \alpha_k p_k$
4. $d_{k+1} := d_k - \alpha_k w_k$
5. $f_k := A^T w_k$
6. $r_{k+1} := r_k - \alpha_k(f_k + \mu p_k) \leftarrow$
7. $\beta_k := \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$
8. $p_{k+1} := r_{k+1} + \beta_k p_k$
9. $g_k := \frac{V_k^T w_k + \mu \tilde{Q}_k^T p_k}{w_k^T w_k + \mu p_k^T p_k} \leftarrow$
10. $\tilde{Q}_{k+1} := \tilde{Q}_k - p_k g_k^T$
11. $V_{k+1} := V_k - w_k g_k^T$
12. Solve the minimization problem for $\tilde{y}_{k+1} \leftarrow$
13. $\tilde{x}_{k+1} := x_{k+1} + \tilde{Q}_{k+1}\tilde{y}_{k+1}$
14. $\tilde{d}_{k+1} := d_{k+1} - V_{k+1}\tilde{y}_{k+1}$
15. $S_{k+1} := S_k - f_k g_k^T$
16. $\tilde{r}_{k+1} := r_{k+1} - (S_{k+1} + \mu \tilde{Q}_{k+1})\tilde{y}_{k+1} \leftarrow$

Note: For this algorithm the norm of the discrepancy is not a decreasing function of k.

Numerical examples.

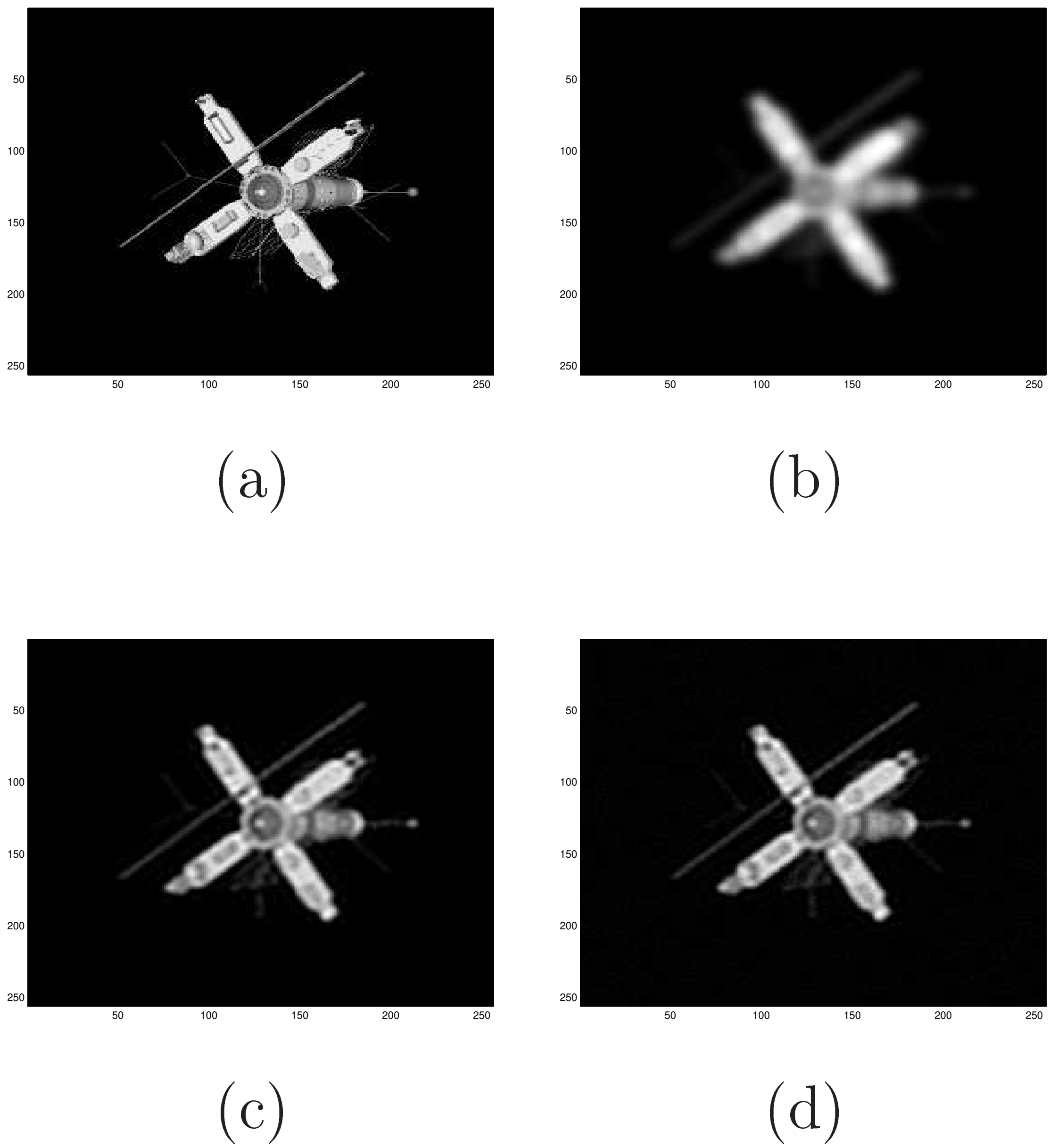
Example 1: 50×50 8-bit image. Gaussian blur with $\sigma = 1.5, \rho = 12\sigma$. Thus A is numerically singular. Noise level 10^{-3}.
$\mathcal{Q} = \{b^\delta\}$.



(a) Exact; (b) Contaminated (c) Restored with 145 CGNR iterations (d) Restored with 53 enriched CGNR steps.

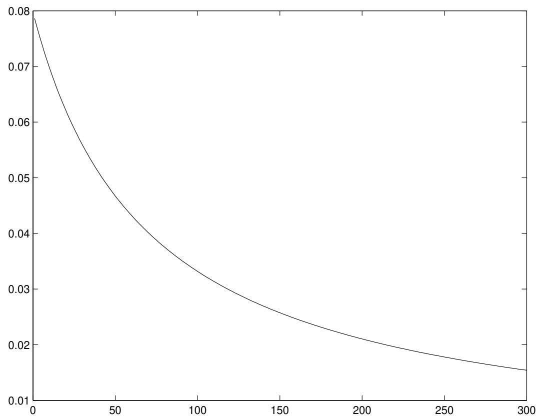
Example 2: Satellite, 256×256; Gaussian blur with $\sigma = 3.5$ and $rho = 12\sigma$. Noise level 10^{-3}.

$\mathcal{Q} = \{b^\delta\}$.

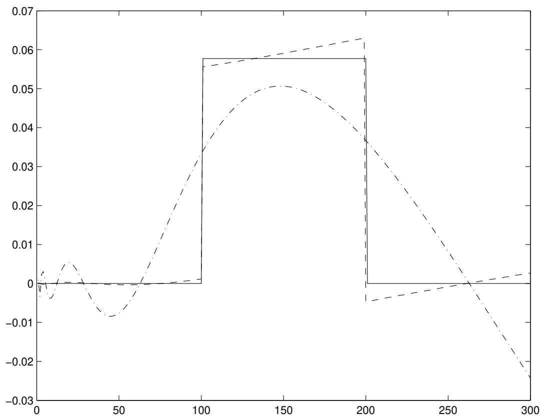


(a) Exact; (b) Corrupted (c) After 238 CGNR iterations
(d) After 81 enriched GCNR iterations.

Example 3: $A = [\frac{1}{i+0.5j}]$; Noise level 10^{-4}
$\mathcal{Q} = \{q_1, q_2, b^\delta\}$.



(a)



(b)

(a) Right-hand side (b) Exact (solid), after 13 steps of CGNR (dash-dotted), after 3 steps of enriched CGNR (dashed).