Computational inverse problems

Nuutti Hyvönen and Juha-Pekka Puska

nuutti.hyvonen@aalto.fi, juha-pekka.puska@aalto.fi

Fourth lecture, March 10, 2021.

2.4 Regularization by truncated iterative methods

For simplicity, in the rest of Chapter 2 we will only consider the case when

$$Ax = y$$

is a system of linear equations, i.e., $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

In the literature there are lots of iterative methods for solving this kind of matrix equations. By “iterative” we mean a method that attempts to solve the problem by finding successive approximations for the solution, starting from some initial guess. Typically, computation of such iterations involves multiplications by A and its adjoint, but not explicit computation of inverse operators. (The *Gaussian elimination* is an example of the opposite: it is a direct, i.e., non-iterative, method that tries to come up with a solution in a finite number of steps.)

Iterative methods are sometimes the only feasible choice if the problem involves a large number of variables (sometimes of the order of millions), making direct methods prohibitively expensive. Iterations are especially practical if multiplications by A are cheap. This is the case, e.g., when A is a multi-diagonal matrix originating from a difference or element approximation for some boundary value problem for an elliptic partial differential operator. (There exist lots of other examples, as well.)

Although iterative solvers have not usually been designed for ill-posed equations, they often posses regularizing properties: If the iterations are terminated before "the solution starts to fit to noise", one often obtains reasonable solutions for inverse problems.

2.4.1 Landweber–Fridman iteration

Banach fixed point iteration

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a vector-valued function. We say that $S \subset \mathbb{R}^n$ is an invariant set for T if

$$T(S) \subset S, \quad \text{i.e.,} \quad T(x) \in S \quad \text{for all } x \in S.$$

Moreover, T is a contraction on an invariant set S if there exists $0 \leq \kappa < 1$ such that

$$\|T(x) - T(y)\| \; < \; \kappa\|x - y\| \qquad \text{for all } x, y \in S.$$

Finally, a vector $x \in \mathbb{R}^n$ is called a fixed point of T if

$$T(x) \; = \; x.$$

.

Theorem. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a contraction on the closed invariant set S. Then there exists a unique fixed point $x \in S$ of T. Furthermore, this fixed point can be found by the following fixed point iteration:*

$$x = \lim_{k \to \infty} x_k, \qquad \text{where } x_{k+1} = T(x_k),$$

for any $x_0 \in S$.

Proof. The proof — although not very complicated — is omitted.

A simple example: Consider the function $T : x \mapsto x^2$ from $\mathbb{R}$ to itself.

(i) Let $S = [0, 1/3]$. Clearly, $T(S) = [0, 1/9] \subset S$ and

$$|T(x) - T(y)| \;=\; |x^2 - y^2| \;=\; |x + y||x - y| \;\leq\; 2/3|x - y|.$$

Hence, there is a unique fixed point, which is given by $\lim x_0^{2^k} = 0$ for every $x_0 \in S$.

(ii) If $S = (0, 1/3]$, the fixed point does not anymore lie in S.

(iii) If $S = [0, 1]$, $T(S) = S$, but T is no longer a contraction:

$$|T(3/4) - T(1/2)| \;=\; 5/16 \;>\; 1/4 = |3/4 - 1/2|.$$

In this case there are two fixed points: $T(0) = 0$ and $T(1) = 1$.

(iv) If, e.g., $S = [0, 5/6]$, there is a unique fixed point $0 \in S$, but its existence is not predicted by the fixed point theorem since T is not a contraction on S.

Landweber–Fridman scheme

Instead of the original equation

$$Ax = y,$$

we will consider the normal equation

$$A^{\mathrm{T}}Ax = A^{\mathrm{T}}y.$$

According to 3. exercise of 1. session, $x \in \mathbb{R}^n$ satisfies the normal equation of and only if it minimizes the residual

$$\|Ax - y\|.$$

Moreover, there exist a unique element of $\mathbb{R}^n$, given by $x^\dagger := A^\dagger y$, that solves the normal equation and is orthogonal to $\mathrm{Ker}(A)$.

(Bear in mind, however, that the use of the pseudoinverse $A^\dagger$ is suspect if the matrix is ill-conditioned, i.e., if $\lambda_1/\lambda_p \gg 1$, where $p = \mathrm{rank}(A)$.)

We define an affine mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(x) \;=\; x + \beta(A^{\mathrm{T}}y - A^{\mathrm{T}}Ax), \qquad \beta \in \mathbb{R}.$$

Notice that any solution of the normal equation is a fixed point of T. We will show that if β is small enough there is only one fixed point of T in $\mathrm{Ker}(A)^{\perp}$, namely $x^{\dagger}$, and it can be reached by the fixed point iteration if $x_0 = 0$.

Theorem. *Let $0 < \beta < 2/\lambda_1^2$ be fixed. Then, the fixed point iteration*

$$x_{k+1} \;=\; T(x_k), \qquad x_0 = 0,$$

converges towards $x^{\dagger}$ as $k \to \infty$.

Proof. Set $S = \mathrm{Ker}(A)^{\perp} = \mathrm{Ran}(A^{\mathrm{T}})$. Clearly, $T(S) \subset S$ since

$$T(x) \;=\; x + A^{\mathrm{T}}\,(\beta y - \beta A x) \in \mathrm{Ran}(A^{\mathrm{T}})$$

for all $x \in \mathrm{Ran}(A^{\mathrm{T}})$. Thus, S is invariant under T.

Recall that A and its transpose can be represented with the help of A's singular system as

$$Ax \;=\; \sum_{j=1}^{p} \lambda_j (v_j^{\mathrm{T}} x) u_j \qquad \text{and} \qquad A^{\mathrm{T}} y \;=\; \sum_{j=1}^{p} \lambda_j (u_j^{\mathrm{T}} y) v_j,$$

where $p = \mathrm{rank}(A)$ and λ_j are the positive singular values of A. The orthonormal sets of vectors $\{v_j\}_{j=1}^{p}$ and $\{u_j\}_{j=1}^{p}$ span $S = \mathrm{Ker}(A)^{\perp}$ and $\mathrm{Ran}(A)$, respectively. In particular,

$$x \;=\; \sum_{j=1}^{p} (v_j^{\mathrm{T}} x) v_j \qquad \text{for all } x \in S.$$

Let $x, z \in S$ and note that also $x - z \in S$. We have

$$\begin{aligned} T(x) - T(z) &= (x - z) - \beta A^{\mathrm{T}} A (x - z) \\ &= \sum_{j=1}^{p} (v_j^{\mathrm{T}}(x - z)) v_j - \beta \sum_{j=1}^{p} \lambda_j^2 (v_j^{\mathrm{T}}(x - z)) v_j \\ &= \sum_{j=1}^{p} (1 - \beta\lambda_j^2)(v_j^{\mathrm{T}}(x - z)) v_j. \end{aligned}$$

As λ_1 is the largest of the singular values, it holds by assumption that

$$-1 \;<\; \beta\lambda_j^2 - 1 \;\le\; \beta\lambda_1^2 - 1 \;<\; 2 - 1 \;=\; 1, \qquad \text{for all } j = 1, \dots, p.$$

Hence, we see that

$$\kappa \;:=\; \max_{j=1,\dots,p} |\beta\lambda_j^2 - 1| \;<\; 1.$$

In consequence,

$$\begin{aligned}\|T(x)-T(z)\|^2 &\leq \sum_{j=1}^{p}(1-\beta\lambda_j^2)^2(v_j^{\mathrm{T}}(x-z))^2\\ &\leq \kappa^2\sum_{j=1}^{p}(v_j^{\mathrm{T}}(x-z))^2 \ = \ \kappa^2\|x-z\|^2,\end{aligned}$$

which shows that T is a contraction on S. As S is also a closed invariant set for T, we know that there exists a unique fixed point of T in S.

To complete the proof, we recall that $x^\dagger = A^\dagger y$ belongs to $S = \mathrm{Ker}(A)^\perp$ and satisfies the normal equation (see exercise 3. of session 1.). Furthermore, since $x_0 = 0$ is in S — it is orthogonal to all vectors —, the fixed point iteration starting from x_0 converges to $x^\dagger$. □

Regularization properties of Landweber–Fridman

From now on we will assume that $0 < \beta < 2/\lambda_1^2$.

In the third exercise session, it will be shown that the kth iterate of the Landweber–Fridman iteration can be written explicitly:

$$x_k = \sum_{j=1}^{p} \frac{1}{\lambda_j}\big(1 - (1-\beta\lambda_j^2)^k\big)(u_j^{\mathrm{T}} y)v_j, \qquad k = 0, 1, \ldots . \tag{8}$$

Since $|1 - \beta\lambda_j^2| < 1$ by assumption,

$$(1 - \beta\lambda_j^2)^k \to 0 \qquad \text{as } k \to \infty,$$

which is what one would expect since

$$x^\dagger = \sum_{j=1}^{p} \frac{1}{\lambda_j}(u_j^{\mathrm{T}} y)v_j.$$

However, while $k \in \mathbb{N}$ is finite, the coefficients of the terms $(u_j^{\mathrm{T}} y)v_j$ appearing in the series representation (8) satisfy

$$\begin{aligned}
\frac{1}{\lambda_j}\big(1-(1-\beta\lambda_j^2)^k\big) &= \frac{1}{\lambda_j}\left(1-\sum_{l=0}^{k}\binom{k}{l}(-1)^l\beta^l\lambda_j^{2l}\right) \\
&= \frac{1}{\lambda_j}\sum_{l=1}^{k}\binom{k}{l}(-1)^{l+1}\beta^l\lambda_j^{2l} \\
&= \sum_{l=1}^{k}\binom{k}{l}(-1)^{l+1}\beta^l\lambda_j^{2l-1},
\end{aligned}$$

which converges to zero as $\lambda_j \to 0$ (for a fixed k).

As a consequence, while k is ‘small enough’, no coefficient of $(u_j^{\mathrm{T}} y)v_j$ in (8) is so large that the component of the measurement noise in the direction u_j is amplified in an uncontrolled manner. (Recall that the corresponding coefficients for Tikhonov regularization are $\lambda_j/(\lambda_j^2+\delta)$.)

Discrepancy principle for Landweber–Fridman

Let the measurement $y \in \mathbb{R}^m$ be a noisy version of some underlying 'exact' data vector $y_0 \in \mathbb{R}^m$, and assume that

$$\|y - y_0\| \approx \epsilon > 0.$$

The Morozov discrepancy principle works for the Landweber–Fridman iteration in approximately the same way as for the truncated SVD and the Tikhonov regularization: Choose the smallest $k \geq 0$ such that the residual satisfies

$$\|y - Ax_k\| \leq \epsilon.$$

Such a stopping rule exists if

$$\epsilon > \|y - Py\| = \|y - A(A^\dagger y)\|,$$

where $P = AA^\dagger$ (see 1. ses., 2. ex.) is the orthogonal projection onto the range of A. Indeed, since the sequence $\{x_k\}_{k=0}^\infty$ converges to $x^\dagger = A^\dagger y$, for any $\epsilon > \|y - Ax^\dagger\|$ there exists $k = k_\epsilon \in \mathbb{N}$ such that

$$\|x_k - x^\dagger\| \leq \frac{1}{\|A\|}(\epsilon - \|y - Ax^\dagger\|),$$

and thus by the reverse triangle inequality,

$$\begin{array}{rcl} \|y - Ax_k\| - \|y - Ax^\dagger\| & \leq & \|(y - Ax_k) - (y - Ax^\dagger)\| \\ & \leq & \|A\|\|x_k - x^\dagger\| \\ & \leq & \epsilon - \|y - Ax^\dagger\|, \end{array}$$

which just means that $\|y - Ax_k\| \leq \epsilon$.

An example: Heat distribution in a rod (revisited)

Recall again the discretized inverse heat conduction problem that was discussed during the second and third lectures. Let `w` be the simulated heat distribution at `T=0.1` with the 'wedge function' as the initial data, and `A` the corresponding propagation matrix `A=expm(TB)`. We add again the same small amount of noise to the measurement:

```
wn = w + 0.001*randn(N-1,1);
```

and use the Morozov discrepancy principle with

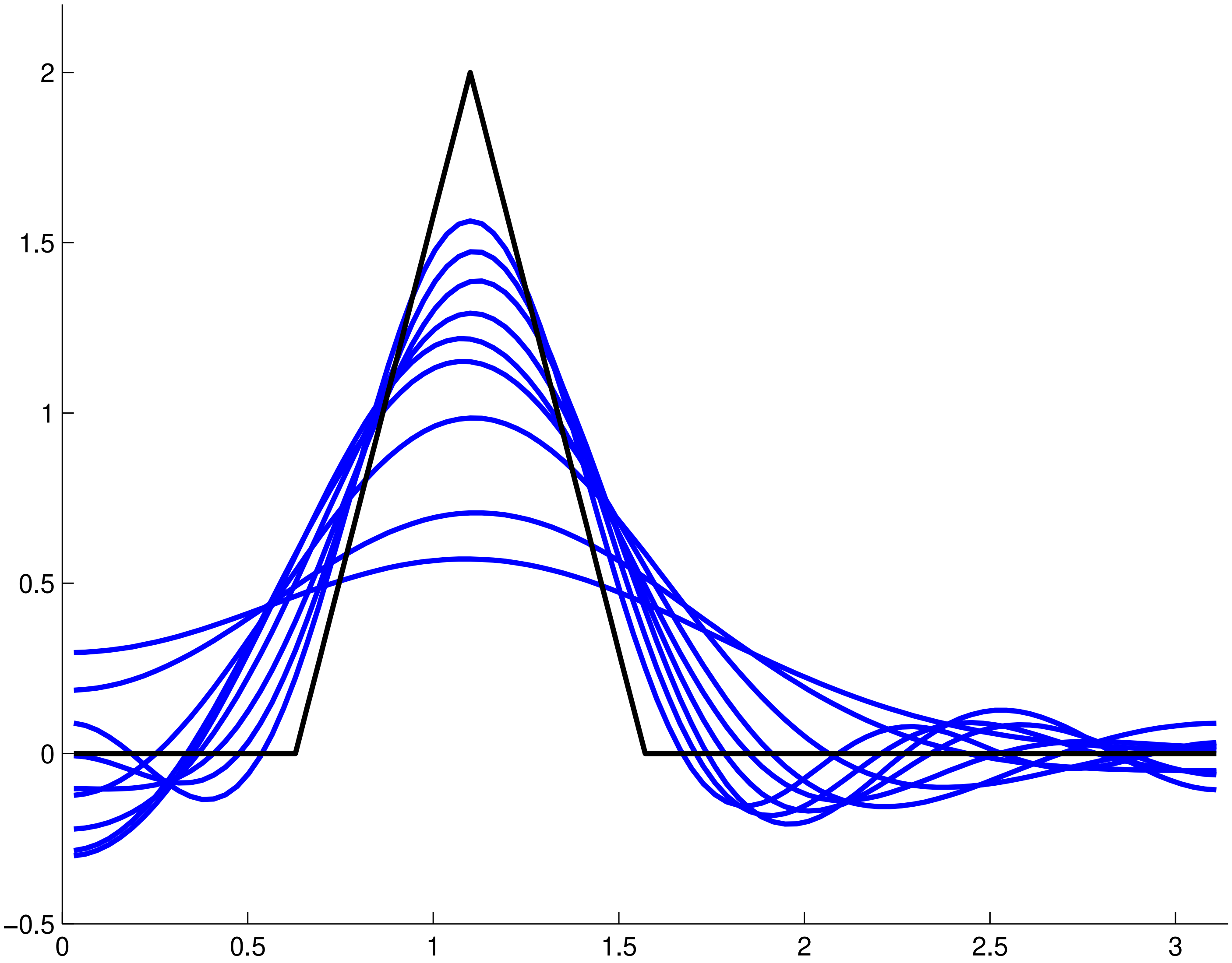
$$\epsilon = \sqrt{99 \cdot 0.001^2} \approx 9.95 \cdot 10^{-3}.$$

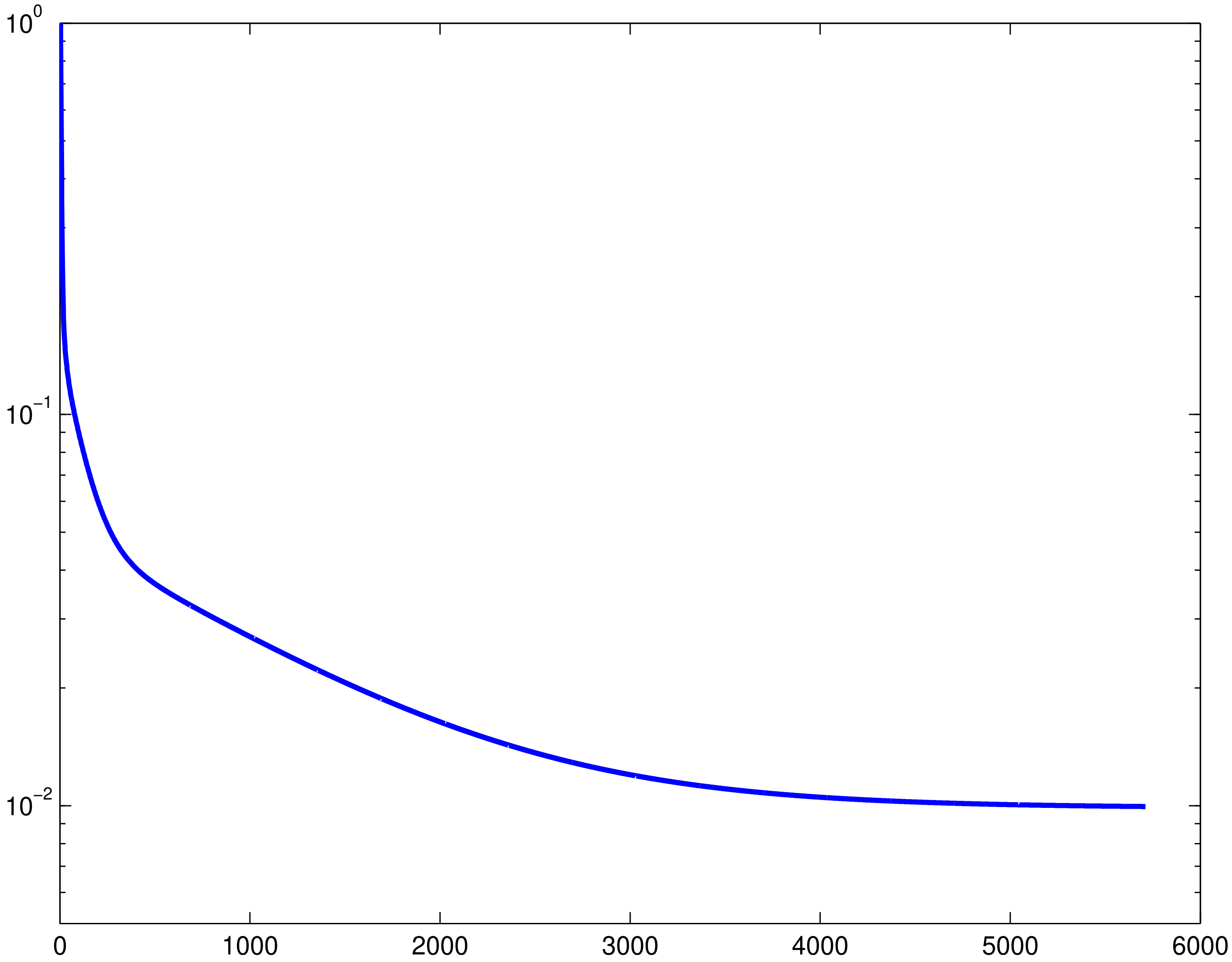
Because the largest singular value of the solution operator $E_T : L^2(0,\pi) \to L^2(0,\pi)$ in the corresponding infinite-dimensional case is 1, it is reasonable to anticipate that the same is also approximately true for A. Thus, we choose $\beta = 1 < 2/1 \approx 2/\lambda_1^2$.

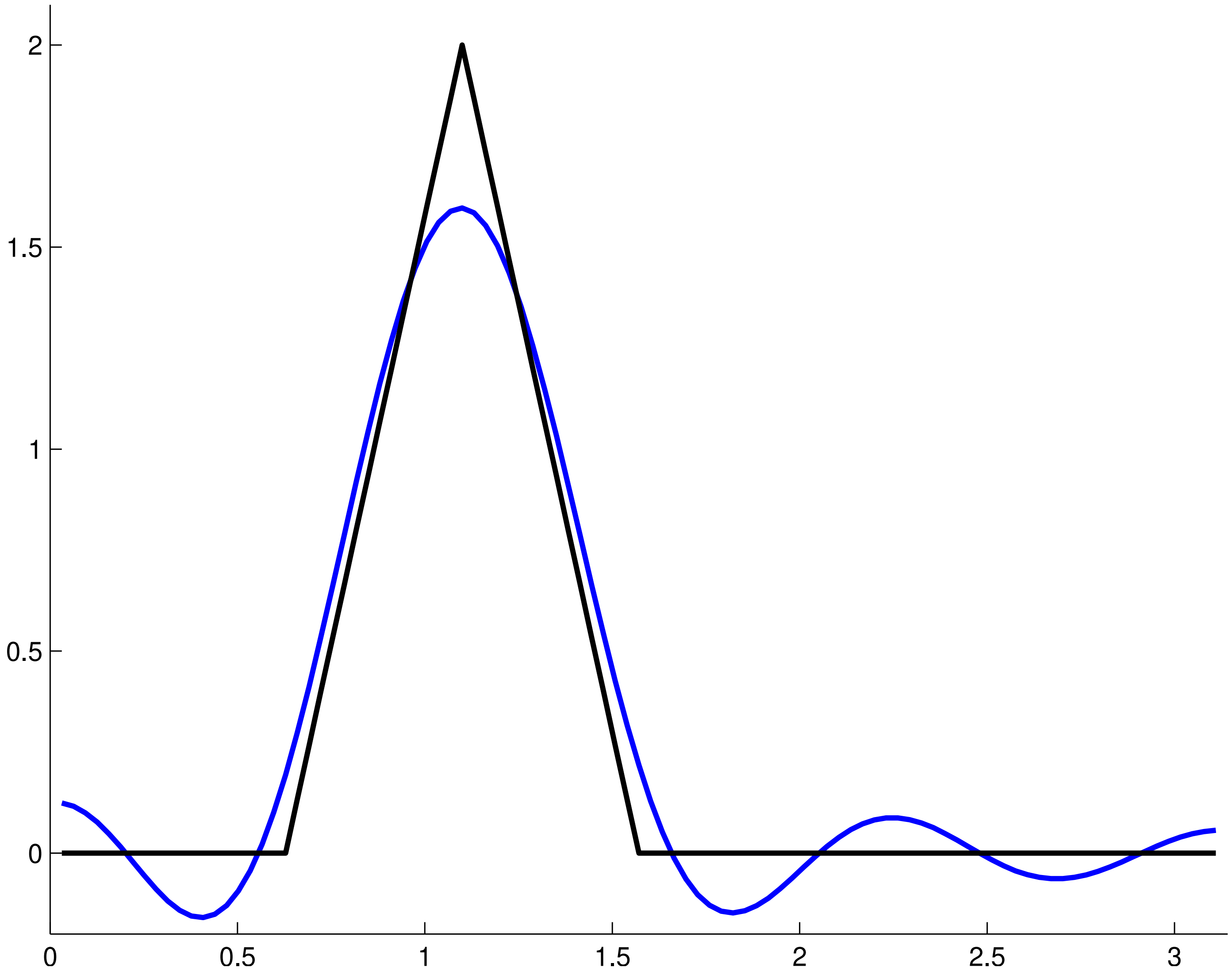
The implementation of the Landweber–Fridman iteration with the Morozov discrepancy principle in Matlab is straightforward. Bear in mind, however, that matrix-matrix products are far more expensive to compute than matrix-vector products. Hence, you should either compute and store the product $A^{\mathrm{T}}A$ before you start iterating or use parentheses to avoid computing this product during the iteration:

```
flw = flw + beta*(A'*wn - A'*(A*flw));
```

With the particular realization of the measurement noise, the Morozov discrepancy principle was satisfied by the iterate corresponding to $k = 5712$. In the following, we visualize the evolution of the Landweber–Fridman iteration for $k = 1, 2, 7, 20, 54, 148, 403, 1096, 2980$, show the residual as a function of k, and plot the solution corresponding to the discrepancy principle.


2
1.5
1
0.5
0
−0.5
0
0.5
1
1.5
2
2.5
3


10^0
10^−1
10^−2
0
1000
2000
3000
4000
5000
6000


2
1.5
1
0.5
0
0
0.5
1
1.5
2
2.5
3