

## Ill-Posed Problems

- Consider the Fredholm integral equation of the first kind:

$$b(t) = \int_0^1 K(s, t)x(t) dt, \quad 0 \leq t \leq 1,$$

where  $K$  is square integrable.

- By Schmidt's theory this equation is equivalent to an infinite matrix equation

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} \kappa_1 & 0 & 0 & \cdots \\ 0 & \kappa_2 & 0 & \cdots \\ 0 & 0 & \kappa_3 & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \vdots \end{pmatrix}$$

Where

$$1. \kappa_1 \geq \kappa_2 \geq \kappa_3 \geq \cdots,$$

$$2. \sum \kappa_i^2, \sum \xi_i^2, \sum \beta_i^2 < \infty.$$

- Condition 2 implies that  $\kappa_i \rightarrow 0$ .
- For a solution to exist the  $\beta_i$  must satisfy the Picard condition

$$\sum \frac{\beta_i^2}{\kappa_i^2} < \infty$$

## Enter Error

- Many physical phenomena are accurately modeled by Fredholm equations.
- In practice the output will be contaminated with error; that is, instead of  $b$  we will observe  $b + e$ .
- The components  $\epsilon_i$  of  $e_i$  will generally be of a size, and therefore  $e$  will not satisfy the Picard condition.

$$\hat{\xi}_i = \frac{\xi_i + \epsilon_i}{\kappa_i} \rightarrow \infty$$

- In many instances the  $\xi_i$  for large  $i$  correspond to highly oscillatory, physically meaningless components of the solution and can be ignored.
- The process by which this is accomplished is called regularization.

Regularization is the art of ignorance.

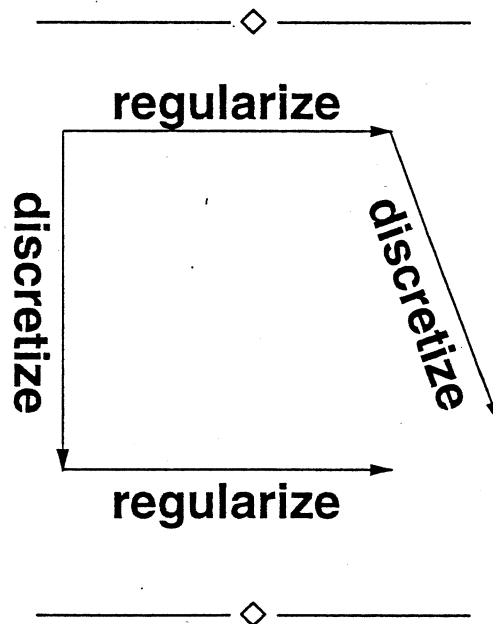
## Discretization

- To solve the problem it must be moved to a finite dimensional space—a process called discretization.
- There are two ways to combine this with regularization.

Discretize then regularize

Regularize then discretize

- They need not be the same.



## SVD Regularization

- SVD regularization truncates the SVD so that the errors in the later component never have a chance to contaminate the solution. Specifically, one solves the system

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \kappa_1 & 0 & \cdots & 0 \\ 0 & \kappa_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \kappa_k \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{pmatrix}$$

- The integer  $k$  is the *regularization parameter* for the procedure.
- The square error is

$$\sum_{i=1}^k \frac{\epsilon_i^2}{\kappa_i^2} + \sum_{i=k+1}^{\infty} \xi_i^2.$$

- The first term grows with  $k$  while the second shrinks. The proper choice of  $k$  will minimize the error.

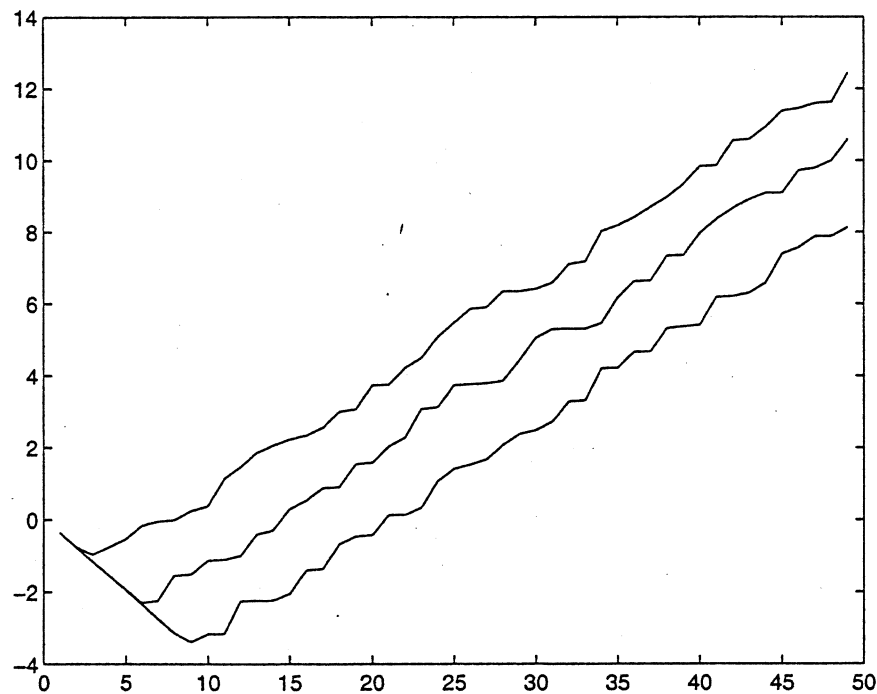
## SVD Regularization (Continued)

- In the following example:

1.  $\kappa_i = 0.5^i$

2.  $\xi_i = 0.4^i$

- The error was  $10^{-2}, 10^{-4}, 10^{-6}$  times a vector of standard normal deviates.



## Tikhonov–Phillips Regularization

- Since large solutions are bad solutions, another regularization strategy is to restrict the size of the solution.
- Since oscillatory solutions are also bad, one could also require the solution to be smooth.
- Tikhonov–Phillips regularization takes as an approximation to  $x$  the solution of the problem

$$\begin{aligned} & \text{minimize } \|b - K\hat{x}\| \\ & \text{subject to } \|L\hat{x}\| \leq \lambda, \end{aligned}$$

where  $L$  is an operator.

- If  $L = I$  then the restriction is on the size of the problem. If  $L$  is a differential (or difference) operator, it smooths.
- For some parameter  $\alpha^2$  The procedure is equivalent to the least squares problem of minimizing

$$\|b - K\hat{x}\|^2 + \alpha^2 \|L\hat{x}\|^2,$$

whose normal equations are

$$(K^*K + \alpha^2 L^*L)\hat{x} = K^*b.$$

## Tikhonov–Phillips Regularization (Continued)

- The regularized solution can be written explicitly in terms of the SVD.

$$\hat{\xi}_i = \frac{\kappa_i^2 \xi_i + \kappa_i \epsilon_i}{\kappa_i^2 + \alpha^2}$$

- There are three regimes.

$$\kappa_i \text{ well below } \alpha \quad \hat{\xi}_i \cong 0$$

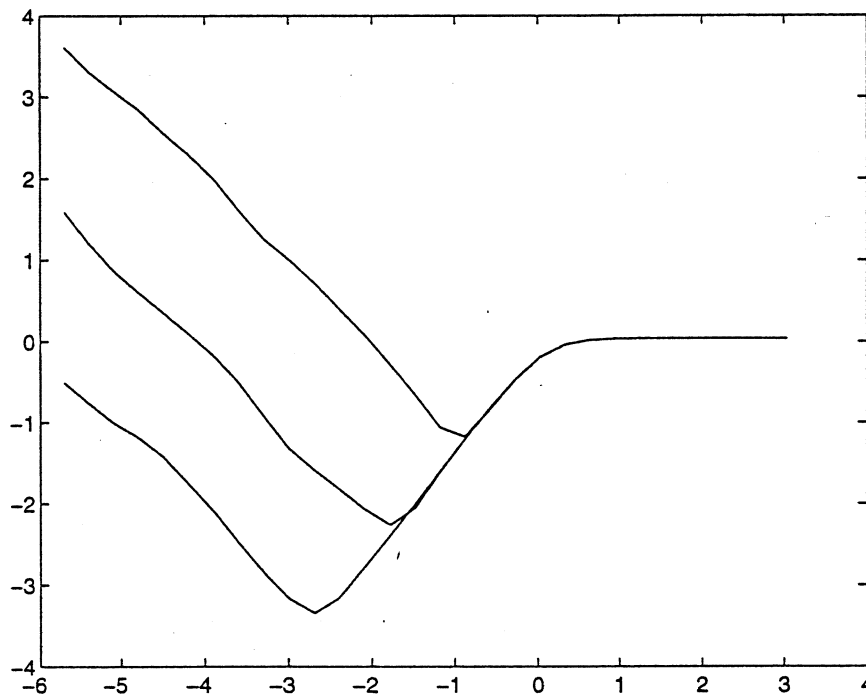
$$\kappa_i \text{ near } \alpha \quad \hat{\xi}_i \cong \frac{1}{2}(\xi_i + \epsilon_i/\kappa_i)$$

$$\kappa_i \text{ well above } \alpha \quad \hat{\xi}_i \cong \xi_i + \epsilon_i/\kappa_i$$

- If  $\alpha$  is too large, we damp out important information.
- If  $\alpha$  is too small, we magnify the error.

## Tikhonov–Phillips Regularization (Continued)

- Here is a plot of the logarithm of  $\alpha$  vs. the logarithm of the error for our example.



- The horizontal asymptote is where the solution is effectively zero.

## Krylov Methods

- A Krylov sequence is a sequence of the form

$$r, Kr, K^2r, K^3r, \dots$$

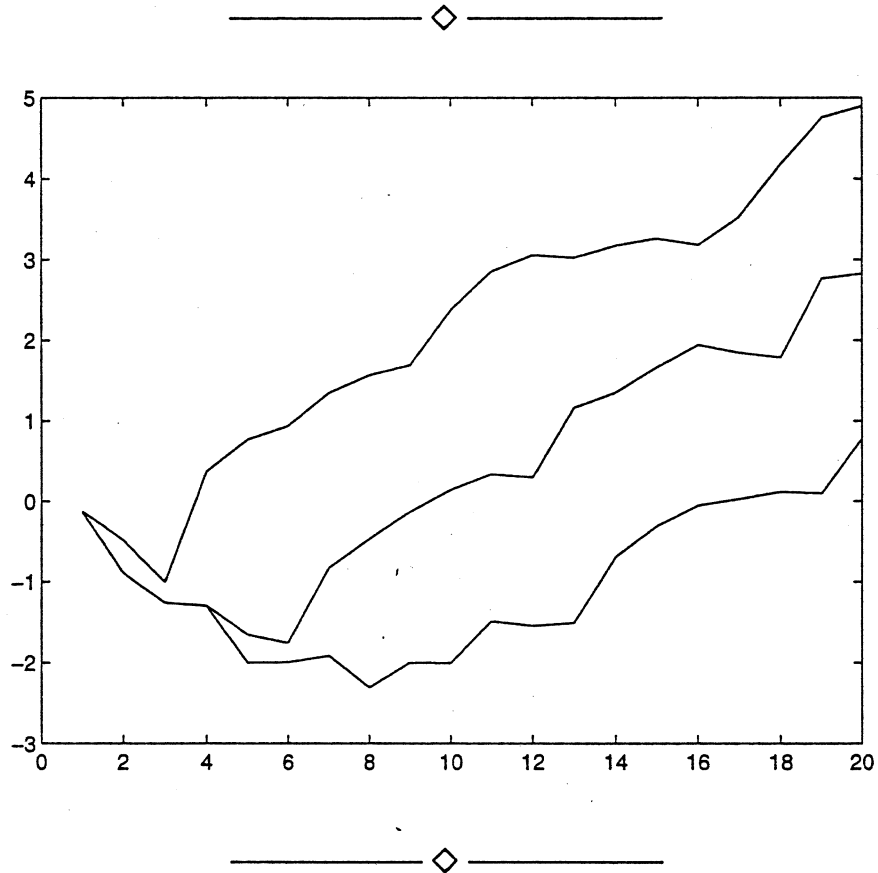
- A Krylov sequence generally contains good approximations to the dominant eigenvectors of a symmetric operator.
- If we take  $K$  as in our example and set  $r = (1, .5, .25, \dots)^T$ , then the accuracy of the best approximation to  $e_1$  is

$i$	error
1	5e-01
2	1e-01
3	2e-02
4	1e-03
5	3e-05

- Thus a good approximation to the dominant part of the solution lies in the Krylov subspace.
- There are several algorithms for extracting a solution — e.g., conjugate gradients, GMRES. Some of them do not require much storage and are suitable for very large problems.

## Krylov Methods (Continued)

The following is a plot of the error as a function of iteration number for GMRES.



- The behavior of these methods as regularizers is not well understood.
- They are quite sensitive to error.

# Krylov Methods (Continued)

