Optimization

Math Lecture 5

Nicholas Dwork

Main Goal

Find x **such that** $Ax \approx b$**.**

Three possibilities

- **There doesn't exist any** x **that satisfies**
- **There exists exactly 1** x **that satisfies**
- **There exists infinitely many** x **that satisfies**

One Solution

If there exists a solution, then we seek the x that satisfies $A\,x = b$.

No Solutions

If there doesn't exist any x **such that** $A\,x = b$ **then we seek the smallest** x **that minimizes**

$$\|A\,x = b\|_2$$

Infinite Solutions

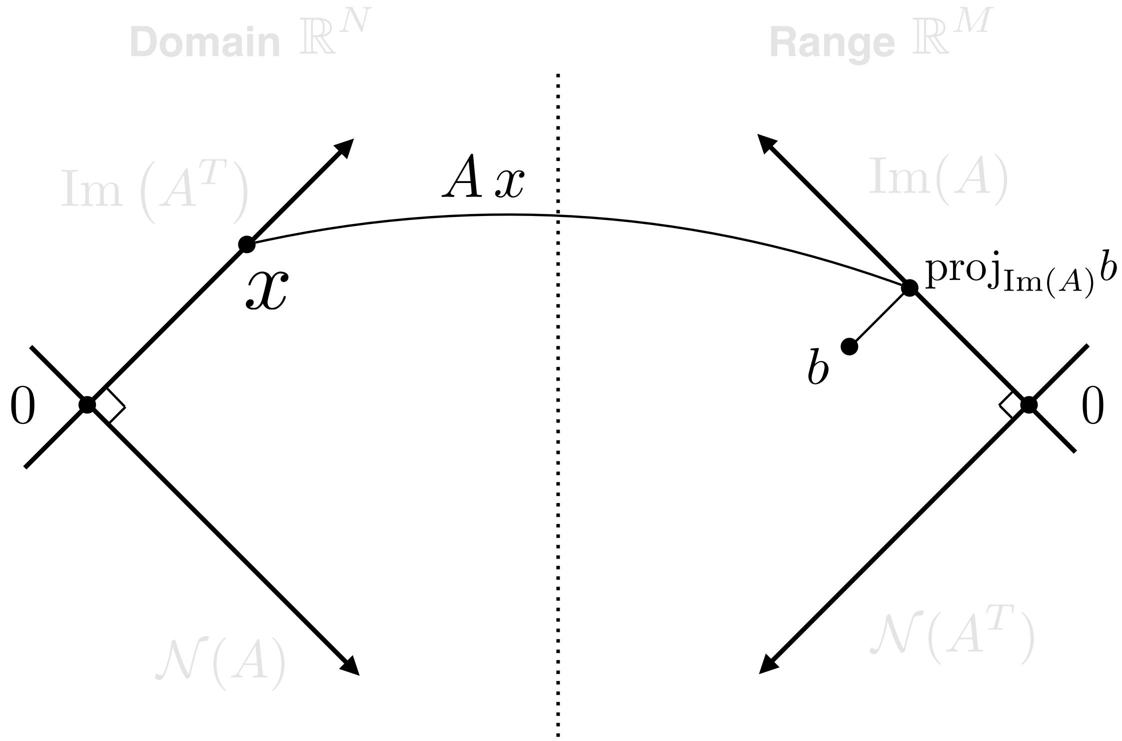
If there exist infinite solutions, then we seek the smallest x that satisfies

$$A\,x = b$$

Pseudo-Inverse

The Pseudo-Inverse of A is the matrix $A^\dagger$ such that $A^\dagger\, b$ is the solution to the following optimization problem (for any b).

$$\begin{array}{rl} \text{minimize} & \|x\|_2 \\ \text{subject to} & Ax = \text{proj}_{\text{Im}(A)} b \end{array}$$



Most common situation is that there are no solutions and A is tall and skinny with linearly independent columns

$$A\,x = b$$

Left multiply by A^T

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

The expression above is called the "Normal Equations"

$$A^T A x = A^T b$$

Since A has linearly independent columns, $A^T A$ is invertible.

$$x = \left(A^T A\right)^{-1} A^T b$$

Now we see how to determine the pseudo-inverse of A for this situation

$$A^{\dagger} = \left(A^T A\right)^{-1} A^T$$

9

Mathematical computer programs have the pseudo-inverse solution built in.

In python

```
x = numpy.linalg.solve( A, b )
```

In Matlab

```
x = A \ b;
```

Regularization

If there are infinite solutions, we must do more to make the solution unique.

This is accomplished through Regularization

$$\text{minimize} \quad \|Ax - b\|_2 + \gamma\, R(x)$$

R is the regularization function.

This is also often done even if there are no solutions to make the problem better behaved.

Tikhonov Regularization

$$\text{minimize} \quad \|A\,x - b\|_2 + \gamma\,\|\Gamma\,x\|_2$$

Γ **is called the Tikhonov matrix**

γ **is the regularization parameter.
Trades off importance of data matching and regularization**

Tikhonov Regularization

$$\text{minimize} \quad \|A\,x - b\|_2 + \gamma\,\|\Gamma\,x\|_2$$

Example Tikhonov matrices:

D **-** *Dx* **is a vector of all the horizontal and vertical differences. Used if** *x* **is expected to be smooth**

I **- identity matrix is used when x is expected to be small**

Tikhonov Regularization

$$\text{minimize} \quad \|A\,x - b\|_2 + \gamma\,\|\Gamma\,x\|_2$$

Can be combined into a minimization of a single term.

You will do this for homework.