Linear Algebraic Equations

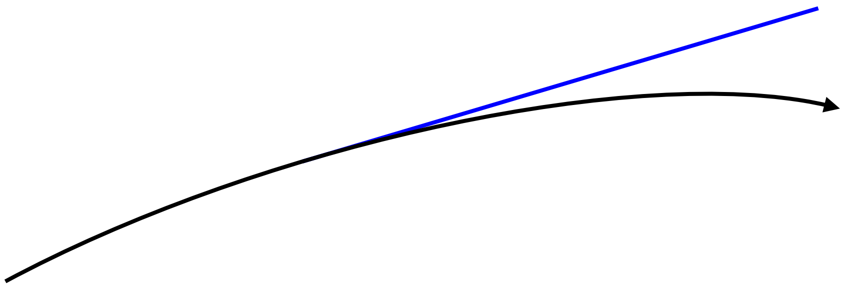
Vasily Arzhanov

Reactor Physics, KTH

Overview

- Vectors, Matrices
- Vector-Matrix Multiplication
- Dot/Inner Product
- Vector Norms
- Induced Matrix Norms
- Condition Number
- Relative Error in Solution
- Solution Improvement

Linearity in Small



$$f(x+\Delta x)=f(x)+f'(x)\Delta x+O\left(\Delta x^2\right)$$

$$-dn(t)=\lambda\cdot n(t)\cdot dt \longrightarrow n(t)=n(0)e^{-\lambda t}$$

Linear System of Equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}$$

Equivalent Linear Systems

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \ldots + \underset{\substack{\times \\ \alpha}}{a_{nn}}x_n = b_n \quad \times\alpha
\end{cases}
\quad \left.\begin{matrix} \\ \\ \end{matrix}\right\} +
$$

Matrices

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}; \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

Vector-Matrix Multiplication

$$\mathbf{Ax} = \mathbf{y} \longrightarrow y_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, 2, \ldots, m$$

$$\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \times \begin{bmatrix} * \\ * \\ * \\ * \\ * \end{bmatrix} = \begin{bmatrix} * \\ * \\ * \end{bmatrix}$$

Left Multiplication

$$\mathbf{x}^T \mathbf{A} = \mathbf{y}^T \longrightarrow y_i = \sum_{j=1}^{m} x_j a_{ji}, \qquad i = 1,2,\ldots,n$$

$$\begin{bmatrix} * & * & * \end{bmatrix} \times \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} = \begin{bmatrix} * & * & * & * & * \end{bmatrix}$$

Dot Product

$$\mathbf{x} \cdot \mathbf{y} \equiv \sum_{i=1}^{n} x_i \overline{y}_i = \mathbf{y}^H \mathbf{x} \qquad \begin{bmatrix} * & * & * \end{bmatrix} \times \begin{bmatrix} * \\ * \\ * \end{bmatrix} = \begin{bmatrix} * \end{bmatrix}$$

Cauchy-Schwarz inequality:

$$\left| \mathbf{x} \cdot \mathbf{y} \right| \leq \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \cdot \sqrt{y_1^2 + y_2^2 + \cdots + y_n^2}$$

Matrix Operations

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \longrightarrow c_{ij} = a_{ij} + b_{ij}$$

$$\mathbf{C} = \lambda \mathbf{A} \longrightarrow c_{ij} = \lambda a_{ij}$$

$$\mathbf{C} = \mathbf{A}\mathbf{B} \longrightarrow c_{ij} = \sum_{k} a_{ik} b_{kj}$$

All Matrices

Not commutative $\mathbf{AB} \neq \mathbf{BA}$

Associative $\mathbf{A}\left(\mathbf{BC}\right) = \left(\mathbf{AB}\right)\mathbf{C} \longrightarrow \mathbf{A}^{n}$

Distributive $\mathbf{A}\left(\mathbf{B}+\mathbf{C}\right) = \mathbf{AB}+\mathbf{AC}$

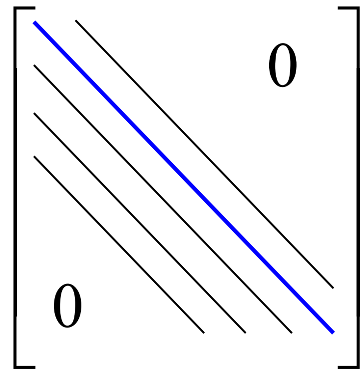
Square Matrices

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \times \begin{bmatrix} * \\ * \\ * \end{bmatrix} = \begin{bmatrix} * \\ * \\ * \end{bmatrix}$$

$$\mathbf{Ax} = \mathbf{0}$$

Special Square Matrices

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{bmatrix}$$

$\mathbf{B} =$ 

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & \ldots & 0 \\ l_{21} & l_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \ldots & l_{mn} \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \ldots & u_{1n} \\ 0 & u_{22} & \ldots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & u_{nn} \end{bmatrix}$$

Some Important Matrices

- Null, $\mathbf{0}$
- Identity, $\mathbf{I}$
- Inverse, $\mathbf{A}^{-1}$, $\quad \mathbf{A} \times \mathbf{A}^{-1} = \mathbf{A}^{-1} \times \mathbf{A} = \mathbf{I}$
- Tridiagonal, $\mathbf{T}$
- Transpose, $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{\mathrm{H}}$
- Positive-definite $\mathbf{x}^{\mathrm{H}}\mathbf{A}\mathbf{x} > 0$ for any $\mathbf{x} \neq \mathbf{0}$.

Over-Determined

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \times \begin{bmatrix} * \\ * \\ * \end{bmatrix} = \begin{bmatrix} * \\ * \\ * \\ * \\ * \end{bmatrix}$$

Under-Determined

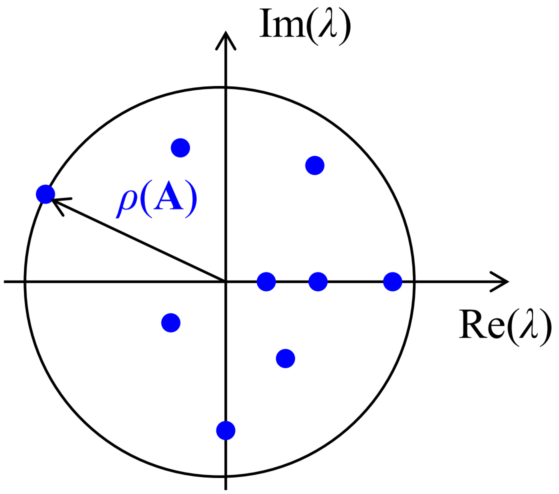
$$\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \times \begin{bmatrix} * \\ * \\ * \\ * \\ * \end{bmatrix} = \begin{bmatrix} * \\ * \\ * \end{bmatrix}$$

Eigenvalues

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

$$\sigma\left(\mathbf{A}\right) \equiv \left\{\lambda_i; i = 1,\ldots,n \;\middle|\; \lambda_i \text{ is eigenvalue}\right\}$$

$$\rho\left(\mathbf{A}\right) \equiv \max_i \left|\lambda_i\right|$$



Vector Norm

$$\mathbf{x}_n \xrightarrow[n\to\infty]{} \mathbf{x}$$

$$\left\|\mathbf{x}\right\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

1) $\left\|\mathbf{x}\right\| > 0 \qquad \forall \mathbf{x} \neq \mathbf{0}$

2) $\left\|\alpha \cdot \mathbf{x}\right\| = \left|\alpha\right| \cdot \left\|\mathbf{x}\right\|$

3) $\left\|\mathbf{x} + \mathbf{y}\right\| \leq \left\|\mathbf{x}\right\| + \left\|\mathbf{y}\right\|$

4) $\left\|\mathbf{x}\right\| = 0 \leftrightarrow \mathbf{x} = \mathbf{0}$

$$\left\|\mathbf{x} - \mathbf{x}_n\right\| < \varepsilon$$

$$\left\|\mathbf{x}\right\|_p \equiv \sqrt[p]{\left|x\right|_1^p + \left|x\right|_2^p + \cdots + \left|x\right|_n^p}$$

Useful Vector Norms

$$\|\mathbf{x}\|_1 \equiv |x_1| + |x_2| + \ldots + |x_n|$$

$$\|\mathbf{x}\|_2 \equiv \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

$$\|\mathbf{x}\|_p \equiv \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p}$$

$$\|\mathbf{x}\|_\infty \equiv \max\left\{|x_1|, |x_2|, \ldots, |x_n|\right\}$$

Generated Norms

$$\left\|\mathbf{x}\right\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

$$\left\|\mathbf{x}\right\|_A = \sqrt{\mathbf{x}^H \mathbf{A}\mathbf{x}}$$

My own norm:

$$\left\|\mathbf{x}\right\|_V \equiv 2\left|x_1\right| + \sqrt{3\left|x_2\right|^2 + \max\left(\left|x_3\right|, 2\left|x_4\right|\right)}$$

Equivalent Norms

$$\|\mathbf{x}\|_\alpha \sim \|\mathbf{x}\|_\beta$$

$$\exists\, C > 0 \;\text{ and }\; D > 0$$

$$C\|\mathbf{x}\|_\alpha \le \|\mathbf{x}\|_\beta \le D\|\mathbf{x}\|_\alpha$$

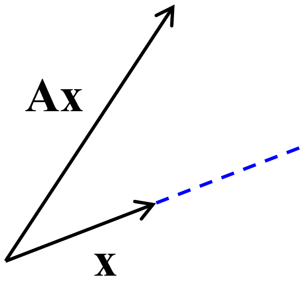
$$\|\mathbf{x}\|_2 \le \|\mathbf{x}\|_1 \le \sqrt{n}\,\|\mathbf{x}\|_2$$

$$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_2 \le \sqrt{n}\,\|\mathbf{x}\|_\infty$$

$$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_1 \le n\|\mathbf{x}\|_\infty$$

Induced Matrix Norm

$$\|\mathbf{A}\| \equiv \max_{\mathbf{x}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \max_{\mathbf{x}} \frac{\|\mathbf{A}\alpha\mathbf{x}\|}{\|\alpha\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$



$$\frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \le \max_{\mathbf{x}} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \|\mathbf{A}\| \longrightarrow \|\mathbf{Ax}\| \le \|\mathbf{A}\| \cdot \|\mathbf{x}\|$$

Statement

In "Numerical Methods for Engineers and Scientists," 2^{nd} edition on P.56, J. D. Hoffman states that

$$\left\| \mathbf{A} \right\|_2 = \min_i \lambda_i \quad \text{(eigenvalue)}$$

Useful Matrix Norms

$$\|\mathbf{A}\|_1 = \max_j \sum_i |a_{ij}|$$

$$\underbrace{\begin{array}{ccc}\downarrow & \downarrow & \downarrow \\ \downarrow & \downarrow & \downarrow \\ \downarrow & \downarrow & \downarrow\end{array}}_{\max}$$

Maximum column sum

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}\left(\mathbf{A}^H\mathbf{A}\right)}$$

$$\|\mathbf{A}\|_\infty = \max_i \sum_j |a_{ij}|$$

$$\left.\begin{array}{ccc}\rightarrow & \rightarrow & \rightarrow \\ \rightarrow & \rightarrow & \rightarrow \\ \rightarrow & \rightarrow & \rightarrow\end{array}\right\}\max$$

Maximum row sum

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$$

Useful Properties

$$\rho(\mathbf{A}) \le \|\mathbf{A}\|$$

$$\|\mathbf{A}^n\|^{1/n} \xrightarrow[n\to\infty]{} \rho(\mathbf{A})$$

$$\|\mathbf{AB}\| \le \|\mathbf{A}\| \cdot \|\mathbf{B}\| \longrightarrow 1 = \|\mathbf{AA}^{-1}\| \le \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$$

$$\|\mathbf{A}\|_2 \le \sqrt{\|\mathbf{A}\|_1 \cdot \|\mathbf{A}\|_\infty}$$

Condition Number

$$\mathbf{b} = \mathbf{A}\mathbf{x} \longrightarrow \|\mathbf{b}\| \le \|\mathbf{A}\| \cdot \|\mathbf{x}\| \longrightarrow \frac{1}{\|\mathbf{x}\|} \le \frac{\|\mathbf{A}\|}{\|\mathbf{b}\|}$$

$$\Delta\mathbf{b} + \mathbf{b} = \mathbf{A}\left(\mathbf{x} + \Delta\mathbf{x}\right)$$

$$\mathbf{A}\Delta\mathbf{x} = \Delta\mathbf{b} \longrightarrow \Delta\mathbf{x} = \mathbf{A}^{-1}\Delta\mathbf{b} \longrightarrow \|\Delta\mathbf{x}\| \le \left\|\mathbf{A}^{-1}\right\| \cdot \|\Delta\mathbf{b}\|$$

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \|\mathbf{A}\| \left\|\mathbf{A}^{-1}\right\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

$$\kappa_\alpha\left(\mathbf{A}\right) \equiv \|\mathbf{A}\|_\alpha \cdot \left\|\mathbf{A}^{-1}\right\|_\alpha \longrightarrow \frac{\|\Delta\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\alpha} \le \kappa_\alpha\left(\mathbf{A}\right) \frac{\|\Delta\mathbf{b}\|_\alpha}{\|\mathbf{b}\|_\alpha}$$

Effect of Perturbations

Perturbation in:

Right hand side, b
$$\frac{\|\Delta \mathbf{x}_b\|}{\|\mathbf{x}\|} \leq \kappa(\mathbf{A}) \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}$$

Matrix, $\mathbf{A}$
$$\frac{\|\Delta \mathbf{x}_A\|}{\|\tilde{\mathbf{x}}\|} \leq \kappa(\mathbf{A}) \frac{\|\Delta \mathbf{A}\|}{\|\mathbf{A}\|} \qquad \frac{\|\Delta \mathbf{A}\|}{\|\mathbf{A}\|} \ll 1$$

Both $\mathbf{b}$ and $\mathbf{A}$
$$\frac{\|\Delta \mathbf{x}\|}{\|\tilde{\mathbf{x}}\|} \leq \frac{\kappa(\mathbf{A})}{1-\kappa(\mathbf{A})\dfrac{\|\Delta \mathbf{A}\|}{\|\mathbf{A}\|}} \left(\frac{\|\Delta \mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|} \right)$$

Interpretation of κ(A)

$$\kappa(\mathbf{A}) \equiv \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\|$$

$$\|\mathbf{A}^{-1}\| = \max_{\mathbf{y}} \|\mathbf{A}^{-1}\mathbf{y}\| / \|\mathbf{y}\| = \max_{\mathbf{y}} \frac{1}{\|\mathbf{y}\| / \|\mathbf{A}^{-1}\mathbf{y}\|} = \frac{1}{\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|}$$

$$\kappa(\mathbf{A}) = \frac{\max_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|}{\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x}\| / \|\mathbf{x}\|}$$

A matrix with large condition number
is said to be ill-conditioned.

Lower Bound of κ(A)

$$\max\left|\lambda_i\right| \le \left\|\mathbf{A}\right\|$$

$$\frac{1}{\min\left|\lambda_i\right|} = \max\left|\lambda_i^{-1}\right| \le \left\|\mathbf{A}^{-1}\right\|$$

$$\frac{\max\left|\lambda_i\right|}{\min\left|\lambda_i\right|} \le \kappa\left(\mathbf{A}\right)$$

Properties of κ(A)

- $1 \leq \kappa(\mathbf{A}) \leq \infty$
- $\kappa(\mathbf{I}) = 1;\ \kappa(\mathbf{S}) = \infty$ (For any singular)
- $\kappa(\alpha\mathbf{A}) = \kappa\ (\mathbf{A})$
- $\kappa(\mathbf{D}) = \max|d_i|/\min|d_i|$
- $\max|\lambda_i|/\min|\ \lambda_i| \leq \kappa(\mathbf{A})$

Residual

$$\mathbf{Ax} = \mathbf{b} \longrightarrow \tilde{\mathbf{x}} \longrightarrow \mathbf{r} \equiv \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}} \longrightarrow \mathbf{A}\tilde{\mathbf{x}} = \mathbf{b} - \mathbf{r}$$

$$\mathbf{e} \equiv \mathbf{x} - \tilde{\mathbf{x}} \longrightarrow \mathbf{Ae} = \mathbf{r}$$

$$\frac{1}{\kappa(\mathbf{A})} \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \le \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \le \kappa(\mathbf{A}) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

Exact arithmetic

$$\frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \le \kappa(A) \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

Solution Improvement

$$\mathbf{Ax} = \mathbf{b} \longrightarrow \tilde{\mathbf{x}} \longrightarrow \mathbf{r} \equiv \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}$$

$$\mathbf{e} \equiv \mathbf{x} - \tilde{\mathbf{x}}$$

$$\mathbf{Ae} = \mathbf{r} \xrightarrow{\;ideal\;} \mathbf{x} = \tilde{\mathbf{x}} + \mathbf{e}$$

$$\downarrow$$

$$\tilde{\mathbf{e}} \rightarrow \tilde{\tilde{\mathbf{x}}} = \tilde{\mathbf{x}} + \tilde{\mathbf{e}} \rightarrow \left\|\mathbf{x} - \tilde{\tilde{\mathbf{x}}}\right\| < \left\|\mathbf{x} - \tilde{\mathbf{x}}\right\|$$

Iterative Refinement

$$\mathbf{Ax} = \mathbf{b} \longrightarrow \mathbf{x}^{(0)} = \tilde{\mathbf{x}}$$

$$k = 0, 1, \ldots$$

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{Ax}^{(k)} \qquad \mathbf{Ae}^{(k)} = \mathbf{r}^{(k)}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{e}^{(k)}$$

Simple Example

$$\begin{bmatrix} 420 & 210 & 140 & 105 \\ 210 & 140 & 105 & 84 \\ 140 & 105 & 84 & 70 \\ 105 & 84 & 70 & 60 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 875 \\ 539 \\ 399 \\ 319 \end{bmatrix}$$

$$\mathbf{x} = [1,1,1,1]^T$$

Iterative Refinement Example in Single Precision

$\mathbf{x}^{(k)}$	x_1	x_2	x_3	x_4
$\mathbf{x}^{(0)}$	0.999988	1.000137	0.999670	1.000215
$\mathbf{x}^{(1)}$	0.999994	1.000069	0.999831	1.000110
$\mathbf{x}^{(2)}$	0.999996	1.000046	0.999891	1.000070
$\mathbf{x}^{(3)}$	0.999993	1.000080	0.999812	1.000121
$\mathbf{x}^{(4)}$	1.000000	1.000006	0.999984	1.000011

Simple Programming Trick

$$r_i = b_i - (\mathbf{Ax})_i$$

$$(\mathbf{Ax})_i = a_{i,0}x_0 + a_{i,1}x_1 + \cdots + a_{i,n-1}x_{n-1}$$

EP = Extended Precision (Binary80, Quadruple)

```
S = 0.0
for j in range(n):
   S += a[i,j]*x[j]
r[i] = b[i] - S
```

```
S = EP(0.0)
for j in range(n):
   S += EP(a[i,j])*EP(x[j])
r[i] = EP(b[i]) - S
```

Important

- Vectors, Matrices
- Vector-Matrix Multiplication
- Dot/Inner Product
- Vector Norms
- Induced Matrix Norms
- Condition Number
- Relative Error in Solution
- Solution Improvement