Basic Concepts

Vasily Arzhanov

Reactor Physics, KTH

Overview

- Vector Space
- Inner Product, Norms
- Well-Posedness
- Categories of Problems
- Conditioning
- Iterative Solution
- Consistency, Stability, Convergence
- Equivalence Theorem

Vector Space

$$\mathbb{V} = \{ f \}$$

1. $\forall f_1, f_2 \in \mathbb{V} \longrightarrow f_1 + f_2 \in \mathbb{V}$

2. $\forall \alpha \ \forall f \in \mathbb{V} \longrightarrow \alpha f \in \mathbb{V}$

A set of vectors, $\{f_1, f_2, \ldots, f_n\}$, is said to be linearly independent if

$\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_n f_n = 0$ implies $\alpha_1 = \alpha_2 = \ldots = a_n = 0$.

Examples of Vector Space

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

$$\Pi_n = \left\{ p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \right\}$$

$$f(x) \in C^n[a,b]$$

Inner Product

$$\langle \cdot,\cdot \rangle : \mathbb{V}\times\mathbb{V} \mapsto \mathbb{C}; \quad \left(\mathbb{V}\times\mathbb{V} \mapsto \mathbb{R}\right).$$

1) $\langle u,v\rangle = \overline{\langle v,u\rangle} \quad \left(\langle u,v\rangle = \langle v,u\rangle\right)$

2) $\langle u+w,v\rangle = \langle u,v\rangle + \langle w,v\rangle$

$$\langle \alpha u,v\rangle = \alpha\langle u,v\rangle$$

3) $\langle u,u\rangle \geq 0 \quad \langle u,u\rangle = 0 \longrightarrow u=0$

Dot Product

$$\mathbf{u} = [u_1, u_2, \ldots, u_n]$$
$$\mathbf{v} = [v_1, v_2, \ldots, v_n]$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$\mathbf{u} \cdot \mathbf{v} \equiv \sum_{i=1}^{n} u_i v_i$$

Vector-Matrix Multiplication

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

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

Vector-Vector Multiplication

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \qquad \mathbf{v}^T = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$$

$$\mathbf{v}^T \mathbf{u} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = v_1 u_1 + v_2 u_2 + v_3 u_3 = \mathbf{v} \cdot \mathbf{u}$$

Examples of Inner Product

Geometric vectors in 3D

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \quad \langle \mathbf{x}, \mathbf{y} \rangle \equiv |\mathbf{x}| \cdot |\mathbf{y}| \cdot \cos\theta$$

Abstract vectors in 3D

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \quad \langle \mathbf{x}, \mathbf{y} \rangle \equiv x_1 y_1 + x_2 y_2 + x_3 y_3$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \mathbf{x}\mathbf{y}^T = \mathbf{x} \cdot \mathbf{y}^T$$

Arbitrary weights, $w_i > 0$

$$\langle \mathbf{x}, \mathbf{y} \rangle \equiv w_1 x_1 y_1 + w_2 x_2 y_2 + w_3 x_3 y_3$$

Inner Product in $\mathbf{C}^n$

$$\mathbf{x}, \mathbf{y} \in \mathbb{C}^n \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle \equiv \mathbf{y}^H \mathbf{x} = x_1 \overline{y}_1 + x_2 \overline{y}_2 + \cdots + x_n \overline{y}_n$$

Inner Product of Functions

$$f(x), g(x) \in C[a,b]$$

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}dx$$

Cauchy-Schwarz Inequality

$$\left|\left\langle \mathbf{u}, \mathbf{v} \right\rangle\right|^2 \leq \left\langle \mathbf{u}, \mathbf{u} \right\rangle \cdot \left\langle \mathbf{v}, \mathbf{v} \right\rangle$$

$$\left|\int_a^b f(x)\overline{g(x)}dx\right|^2 \leq \int_a^b \left|f(x)\right|^2 dx \cdot \int_a^b \left|g(x)\right|^2 dx$$

Related Definitions

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0 \longrightarrow \mathbf{u} \perp \mathbf{v}$$

$$|\mathbf{u}| \equiv \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$$

$$\cos\theta \equiv \frac{\operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle}{|\mathbf{u}| \cdot |\mathbf{v}|}$$

Convergence in Vector Space

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \qquad f_m \in C[a,b]$$

$$\mathbf{x}^{(m)} \xrightarrow[m\to\infty]{} \mathbf{x}; \qquad f_m(x) \xrightarrow[m\to\infty]{} f(x).$$

$$x_k^{(m)} \xrightarrow[m\to\infty]{} x_k; \quad \forall x \quad f_m(x) \xrightarrow[m\to\infty]{} f(x).$$

Norms in Vector Spaces

$$N: \quad \mathbb{V} \mapsto \mathbb{R} \qquad \forall \mathbf{v} \in \mathbb{V} \;\; \exists\, N(\mathbf{v}) \in \mathbb{R}$$

1. $N(\alpha \mathbf{v}) = |\alpha| N(\mathbf{v})$ (Positive homogeneity/scalability)

2. $N(\mathbf{v} + \mathbf{u}) \le N(\mathbf{v}) + N(\mathbf{u})$ (Triangle inequality/subadditivity)

3. $N(\mathbf{v}) = 0 \longrightarrow \mathbf{v} = 0$ (Separability)

$$\forall \mathbf{v} \in \mathbb{V} \quad N(\mathbf{v}) \ge 0$$

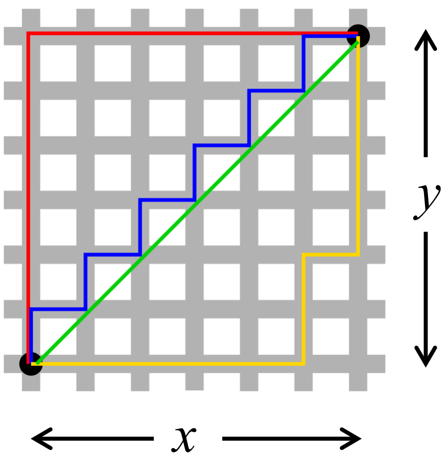
Convergence by Norm

$$\mathbf{x}^{(m)} \xrightarrow[m\to\infty]{} \mathbf{x}; \qquad \left\| \mathbf{x} - \mathbf{x}^{(m)} \right\| \xrightarrow[m\to\infty]{} 0.$$

$$f_m\left(x\right) \xrightarrow[m\to\infty]{} f\left(x\right); \quad \left\| f - f_m \right\| \xrightarrow[m\to\infty]{} 0.$$

Inspiring Examples

Geometric vectors $\mathbf{a} = [x, y, z] \rightarrow \|\mathbf{a}\| = \sqrt{x^2 + y^2 + z^2}$



$$\mathbf{a} = [x, y] \rightarrow \|\mathbf{a}\| = |x| + |y|$$

Manhaten, taxi-cab norm

Useful Vector Norms

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

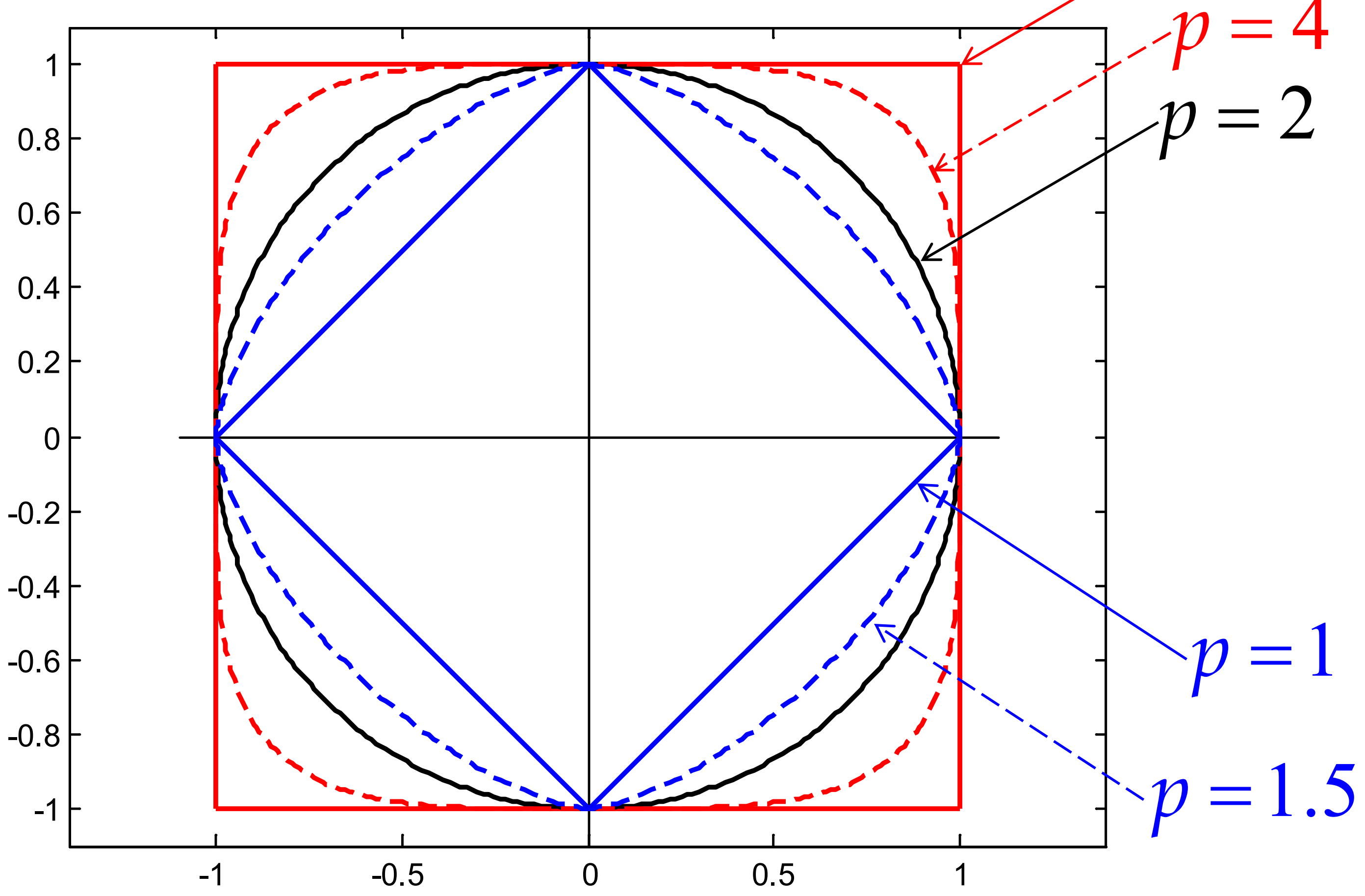
$$\|\mathbf{x}\|_2 \equiv \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

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

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

Visualizing Norms

$$x = [x_1, x_2] \in \mathbb{R}^2 \qquad \|\mathbf{x}\|_p = 1$$



Norms in Functional Spaces

$$\|f\|_1 \equiv \int_a^b |f(x)|\,dx$$

$$\|f\|_2 \equiv \sqrt{\int_a^b |f(x)|^2\,dx} = \sqrt{\langle f, f \rangle}$$

$$\|f\|_p \equiv \left(\int_a^b |f(x)|^p\,dx \right)^{1/p} \xrightarrow[p \to \infty]{} \|f\|_\infty \equiv \max_{x \in [a,b]} |f(x)|$$

Weight

$$w_k > 0; \qquad w(x) > 0$$

$$\|\mathbf{x}\|_p \equiv \left(w_1 |x_1|^p + w_2 |x_2|^p + \cdots + w_n |x_n|^p \right)^{1/p}$$

$$\|f\|_p \equiv \left(\int_a^b w(x) |f(x)|^p \, dx \right)^{1/p}$$

Well-Posedness

$$F(x,d)=0$$

1) A solution exists
2) The solution is unique
3) The solution depends continuously on data

Example of Ill-Posed Problem

$$p(x) = x^4 - (2a-1)x^2 + a(a-1) = 0$$

$$nz = \begin{cases} 0; & a < 0 \\ 2; & 0 \le a < 1 \\ 4; & a \ge 1 \end{cases}$$

Categories of Problems

Find x such that $F(x,d) = 0$

1) Direct problem if F and d are known; x is unknown;
2) Inverse problem if F and x are known; d is unknown;
3) Identification problem if x and d are known; F is unknown

Conditioning

$$F\left(x,d\right)=0; \quad F\left(x+\Delta x, d+\Delta d\right)=0.$$

$$\left\|\Delta x\right\| \le C \cdot \left\|\Delta d\right\|; \qquad K_{abs} = \min C$$

$$\frac{\left\|\Delta x\right\|}{\left\|x\right\|} \le C \cdot \frac{\left\|\Delta d\right\|}{\left\|d\right\|}; \qquad K = \min C$$

Condition Numbers

Absolute $$\|\Delta x\| \le K_{abs} \cdot \|\Delta d\|$$

Relative $$\frac{\|\Delta x\|}{\|x\|} \le K \cdot \frac{\|\Delta d\|}{\|d\|}$$

Infinitesimal Quantities

$$F(x,d)=0; \quad F(x+\delta x, d+\delta d)=0.$$

$$F(x+\delta x, d+\delta d)=F(x,d)+\frac{\partial F}{\partial x}\delta x+\frac{\partial F}{\partial d}\delta d=0$$

$$\delta x=-\frac{\partial F/\partial d}{\partial F/\partial x}\delta d$$

Evaluating Condition Numbers

$$\Delta x \approx -\frac{\partial F/\partial d}{\partial F/\partial x}\Delta d \longrightarrow K_{abs}(d) \approx \left|\frac{\partial F/\partial d}{\partial F/\partial x}\right|$$

$$\frac{\Delta x}{x} \approx -\frac{\partial F/\partial d}{\partial F/\partial x}\frac{d}{x}\frac{\Delta d}{d} \longrightarrow K(d) \approx \left|\frac{\partial F/\partial d}{\partial F/\partial x}\right|\frac{|d|}{|x|}$$

Evaluating Functions

$$f(x+h)-f(x)=f'(\xi)h\approx f'(x)h$$

$$\frac{f(x+h)-f(x)}{f(x)}\approx\frac{f'(x)}{f(x)}h=\frac{xf'(x)}{f(x)}\frac{h}{x}$$

$$x\to fl(x)=x(1+\delta)=x+\underbrace{x\cdot\delta}_{h}\longrightarrow\frac{\Delta f}{f}\approx\frac{xf'(x)}{f(x)}\delta$$

Sensitivity of Simple Roots

$$f(r) = 0; \qquad f'(r) \neq 0; \qquad F(x) = f(x) + \varepsilon g(x) = 0$$

$$f(r+h) + \varepsilon g(r+h) = 0$$

$$f(r) + hf'(r) + \varepsilon g(r) + \varepsilon g'(r)h \approx 0$$

$$h \approx -\varepsilon \frac{g(r)}{f'(r) + \varepsilon g'(r)} \approx -\varepsilon \frac{g(r)}{f'(r)}$$

Example

$$f(x) = (x-1)(x-2)\cdots(x-20) = x^{20} - 210x^{19} + \ldots$$

$$F = (1+\varepsilon)x^{20} - 210x^{19} + \ldots = f + \varepsilon x^{20} = f + \varepsilon g(x)$$

$$h \approx -\varepsilon \frac{g(20)}{f'(20)} = -\varepsilon \frac{20^{20}}{19!} \approx 10^{9}\varepsilon$$

Numerical Instability

$$E_n = \int_0^1 x^n e^{x-1}\,dx = 1 - nE_{n-1}; \qquad E_1 = 1/e.$$

$$E_1 = 0.367879$$

$$E_2 = 0.264242$$

$$\vdots$$

$$E_9 = -0.0684800 \qquad\qquad 9! \approx 3.6\times 10^5$$

Making It Stable

$$E_{n-1} = \frac{1-E_n}{n}; \quad E_n = \int_0^1 x^n e^{x-1}\,dx \le \int_0^1 x^n\,dx = \frac{1}{n+1}$$

$$E_{20} \approx 0$$

$$E_{19} = 0.0500000$$

$$\vdots$$

$$E_9 = 0.0916123 \qquad\qquad 9! \approx 3.6 \times 10^5$$

Iterative Solution

Well-posed: $F\left(x,d\right)=0$

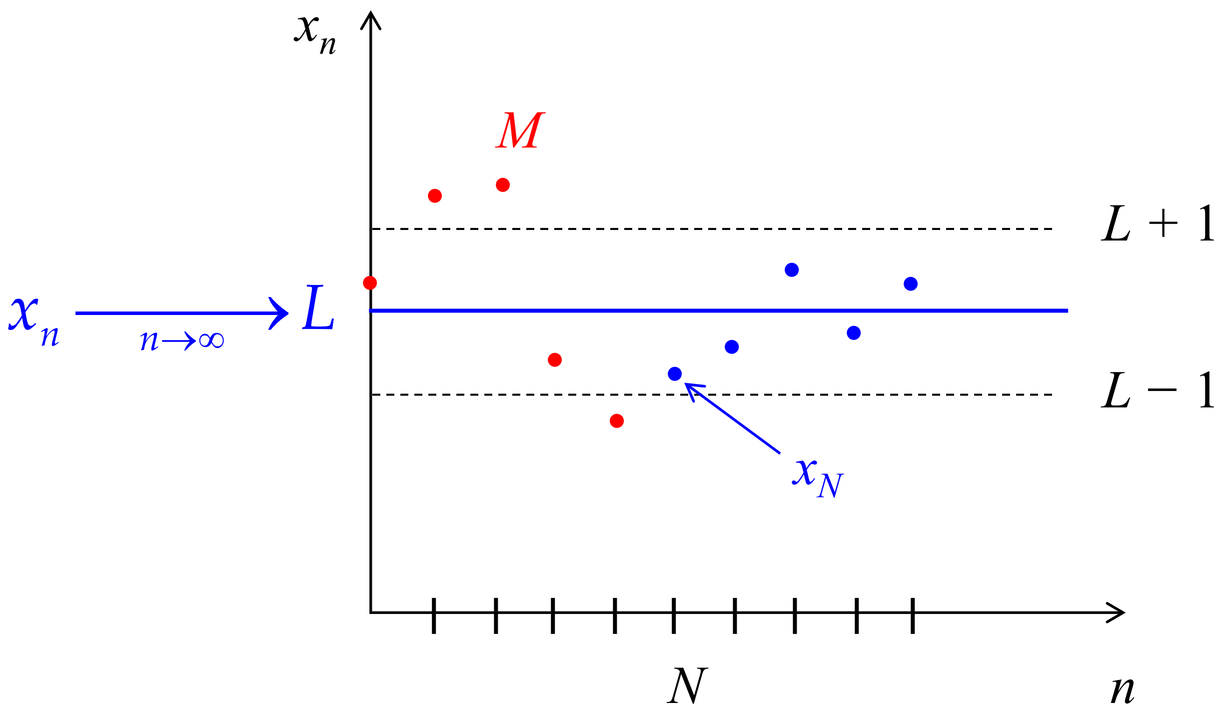
Sequence: $F_n\left(x_n,d_n\right)=0 \qquad \left[F\left(x_n,d_n\right)=0\right]$

Requirement: $d_n \to d; \qquad F_n \to F$

Expect: $x_n \xrightarrow[n\to\infty]{} x$

Convergent Sequences

Any convergent sequence is bounded.



$$|x_n| \le \max(M, L+1)$$

Numerical Stability

$$\forall n \geq 1 \qquad F_n\left(x_n, d_n\right) = 0 \qquad \text{is well-posed}$$

1) There exists a solution, x_n;
2) The solution is unique
3) The solution depends on d_n continuously.

Consequence: x_n is bounded. $\left|x_n\right| \leq C$

Consistency

$$F\left(x, d\right) = 0 \longrightarrow x = x(d)$$

$$F_n\left(x_n, d_n\right) = 0 \qquad \left[F_n\left(x_n, x_{n-1}, \cdots, d_n\right) = 0\right] \qquad x_n = x_n(d)$$

$$F_n\left(x(d), d\right) - F\left(x(d), d\right) = F_n\left(x(d), d\right) \xrightarrow[n \to \infty]{} 0$$

Strongly consistent: $F_n\left(x(d), d\right) = 0$

Example

$$f(x) = 0 \qquad \left[F(x,d) = 0 \right]$$

$$x_0 \longrightarrow x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

$$F_n(x_n, x_{n-1}) \equiv x_n - x_{n-1} + \frac{f(x_{n-1})}{f'(x_{n-1})} = 0$$

Convergence

$$\|x(d) - x_n(d_n)\| \xrightarrow[n\to\infty]{} 0$$

If $F(x,d) = 0$ is well-posed then stability is a necessary condition for convergence.

1) Convergence → Stability
2) Stability → Convergence
   (under certain conditions)

Equivalence Theorem

1) In 1920's, Courant, Friedrichs and Lewy first point out the relationship between stability and convergence;
2) In 1940's, von Neumann identified it more clearly;
3) In 1950's, Lax and Richtmyer brought the issue into an organized form of the equivalence theorem.

For a consistent numerical method, stability is equivalent to convergence.

Kinds of Analysis

$$F\left(x,d\right)=0$$

1) Forward analysis gives bounds on $|x - x^*|$ due to perturbations in data and errors in the numerical method;
2) Backward analysis treats the computed solution as the exact solution of the equation with perturbed data, $F(x^*, d + \Delta d) = 0$;
3) A priori analysis is done prior to computations e.g. by forward or backward analysis;
4) A posteriori analysis evaluates $|x - x^*|$ in terms of the residual, $r = F(x^*, d)$.

Important

- Inner Product, Norms
- Well-Posedness
- Categories of Problems
  - Direct, Inverse, Identification
- Conditioning
- Iterative Solution
- Consistency, Stability, Convergence
- Equivalence Theorem