I.1. TECHNICAL BACKGROUND

Didier HENRION
henrion@laas.fr

Belgian Graduate School on
Systems, Control, Optimization and Networks

Leuven - April and May 2010

Linear algebra

Linear system of equations (LSE)

$$Ax = b$$

with given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ to be found

Can be solved by transforming A into some canonical form
e.g. triangular, diagonal..

Parametrization of all solutions along an affine subspace

$$x = x_0 + Ny$$

with $x_0 \in \mathbb{R}^n$ particular solution satisfying $Ax_0 = b$
and $N \in \mathbb{R}^{n \times (n-r)}$, a null-space basis for A, such that

$$AN = 0, \quad \text{rank}\, A = r$$

Solutions

Either there are no solution

$$r < \text{rank}\,[A \;\; b]$$

or there is a unique solution

$$r = m = n$$

or there is an infinity of solutions

$$r < n$$

Computationally, testing the rank condition, finding a particular solution and parametrizing all solutions can be achieved in one shot with the singular value decomposition, `svd` in Matlab

SVD

Given $A \in \mathbb{R}^{m \times n}$, find U, S, V such that

$$A = USV^T$$

and $U^T U = I_m$, $V^T V = I_n$, $S = \text{diag}\, s_i$, $s_i \geq 0$

Matrices U and V are orthogonal, they respect the geometry of row and column subspaces

Denoting $\bar{x} = V^T x$ and $\bar{b} = U^T b$, our LSE becomes diagonal

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \end{bmatrix}$$

with $S_{11} \in \mathbb{R}^{r \times r}$ the nonzero part of S

Solving LSE via SVD

If $S = S_{11}$ there is a unique solution

$$x = VS^{-1}U^Tb$$

Otherwise if $\bar{b}_2 \neq 0$ there is no solution

Otherwise there is an infinity of solutions

$$x = x_0 + Ny$$

with

$$x_0 = V \begin{bmatrix} S_{11}^{-1}\bar{b}_1 \\ 0 \end{bmatrix}, \quad N = V \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix}$$

QR factorization

More economical than the SVD is the QR factorization

$$A = QR$$

where Q is orthogonal ($Q^T Q = I_m$) and R is upper triangular

The LSE

$$Ax = b$$

is then triangularized

$$Rx = Q^T b$$

and solved by backward subsitution

Similarly we can find a basis for the null-space of A

Geometric interpretation

We have converted an affine subspace in implicit form

$$Ax = b$$

into an affine subspace in explicit or parametric form

$$x = x_0 + Ny$$

Sometimes one formulation is more handy than the other

Later on we will enforce additional constraints

$$x \in K$$

where K is a convex cone (to be defined)

Symmetric matrices

Any $A \in \mathbb{R}^{n \times n}$ has an eigenvalue decomposition

$$AV = VD$$

with D diagonal, and V non-singular (if A is nondefective), both possibly complex-valued, see `eig` in Matlab

If $A = A^T$ is symmetric then V can be orthogonal ($V^T V = I_n$)

$$A = VDV^T$$

with $D = \operatorname{diag} d_i$ real and columns in V can be reordered such that $d_1 \geq d_2 \cdots \geq d_n$, see also `schur` in Matlab

If $A \succeq 0$ we say that A is positive semidefinite and then $d_i \geq 0$

If $A \succ 0$ we say that A is positive definite and then $d_i > 0$, see also `chol` in Matlab

Who is Cholesky ?

André Louis Cholesky (1875-1918) was a French military officer (graduated from Ecole Polytechnique) involved in geodesy

He proposed a new procedure for solving least-squares triangulation problems, just before falling for his country during WWI

I. — NOTICES SCIENTIFIQUES

Commandant BENOIT[1].

NOTE SUR UNE MÉTHODE DE RÉSOLUTION DES ÉQUATIONS NORMALES PROVENANT DE L'APPLICATION DE LA MÉTHODE DES MOINDRES CARRÉS A UN SYSTÈME D'ÉQUATIONS LINÉAIRES EN NOMBRE INFÉRIEUR A CELUI DES INCONNUES. — APPLICATION DE LA MÉTHODE A LA RÉSOLUTION D'UN SYSTÈME DEFINI D'ÉQUATIONS LINÉAIRES.

(Procédé du Commandant CHOLESKY[2].)

Le Commandant d'Artillerie Cholesky, du Service géographique de l'Armée, tué pendant la grande guerre, a imaginé, au cours de recherches sur la compensation des réseaux géodésiques, un procédé très ingénieux de résolution des équations dites *normales*, obtenues par application de la méthode des moindres carrés à des équations linéaires en nombre inférieur à celui des inconnues. Il en a conclu une méthode générale de résolution des équations linéaires.

Inner product

In Euclidean space $\mathbb{R}^n$ the inner product is defined as:

$$< x, y >= x^T y = \sum_i x_i y_i$$

The squared Euclidean norm, or two-norm, of vector x is then

$$\|x\|_2^2 =< x, x >= \sum_i x_i^2$$

In the space of symmetric matrices of size n, isomorphic to $\mathbb{R}^{\frac{n(n+1)}{2}}$, the inner product is defined as:

$$< X, Y >= \text{trace}\,(XY)$$

The squared Frobenius norm of matrix X is then

$$\|X\|_F^2 =< X, X >= \sum_i \sum_j X_{ij}^2$$

Exact vs. approximate

A linear system of equations with rational data has a rational solution = both input data and output data can be represented exactly, or symbolically, on a computer

Some problems cannot be solved exactly, e.g. finding roots of univariate polynomials of degree 5 or more

In this case one must resort to approximate data representation and processing on the computer, using e.g. the IEEE floating point arithmetic

Impact of approximations, or rounding errors, is quantified by numerical analysis

Numerical algorithms

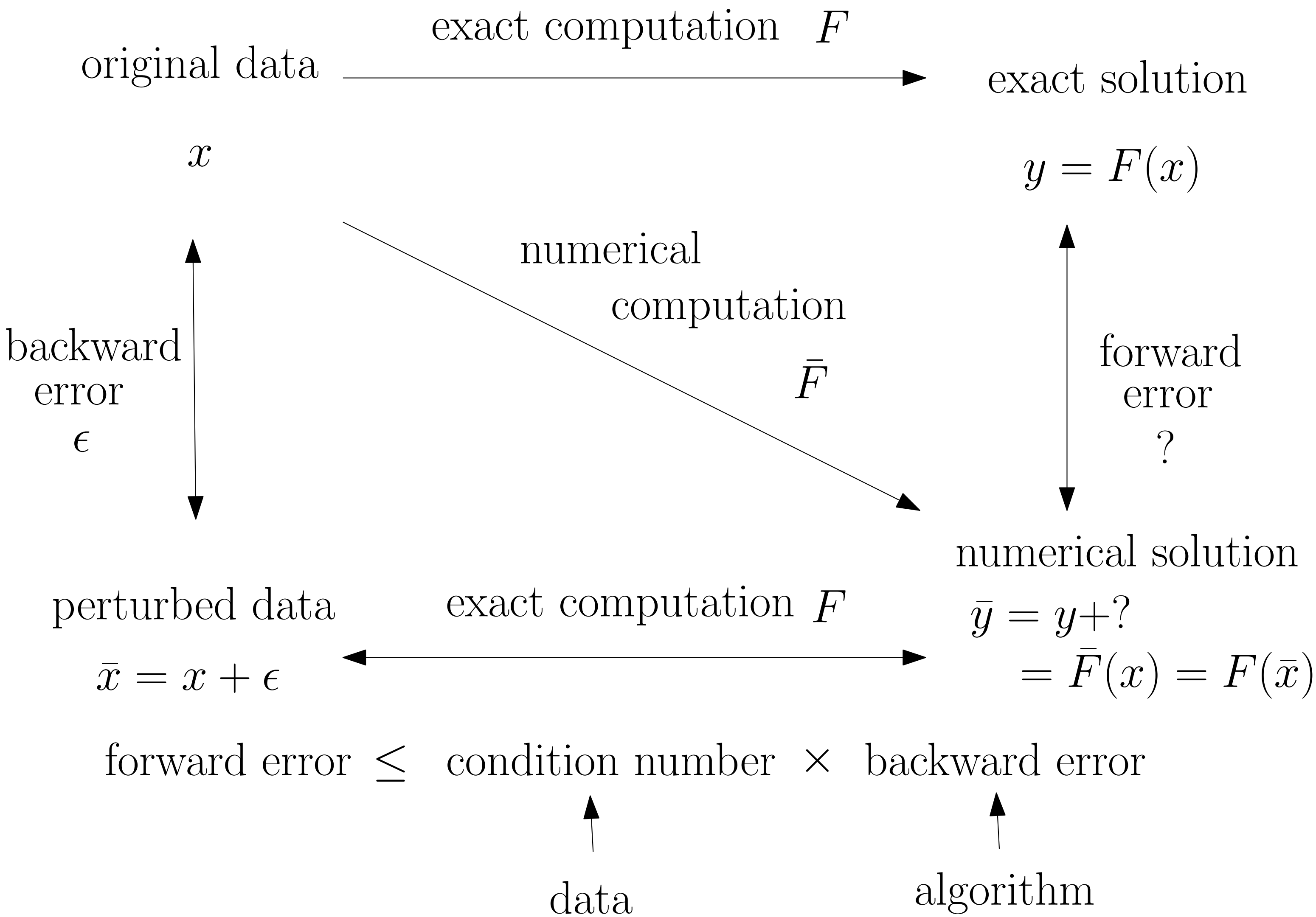
Matlab functions `svd`, `eig`, `schur`, `qr` and `chol` are implementations of algorithms developed by numerical analysts

Core of Matlab = LINPACK, EISPACK (1984) LAPACK (2000)

Key concepts:

- conditioning
- stability
- complexity

J. H. Wilkinson (1960s), N. J. Higham (2002)


original data
x
exact computation F
exact solution
$y = F(x)$
numerical
computation
$\bar{F}$
backward
error
ϵ
forward
error
?
numerical solution
$\bar{y} = y + ?$
$= \bar{F}(x) = F(\bar{x})$
perturbed data
$\bar{x} = x + \epsilon$
exact computation F
forward error $\leq$ condition number $\times$ backward error
data
algorithm

Conditioning

Property of the input data, not of the algorithm

A large condition number means that
the data are sensitive or ill-conditioned

A very-large or infinite condition number corresponds to
an ill-posed problem

Sometimes good conditioning estimates are available

Stability

Property of the algorithm, not of the input data

A small backward error corresponds to
a numerically stable algorithm

A stable algorithm, when applied to well-conditioned data, results in a small forward error, hence the numerical solution is reliable

A stable algorithm, when applied to ill-conditioned data, may generate a large forward error and hence an unreliable numerical solution

Conversely, an unstable algorithm, when applied to well-conditioned data, may also generate a large forward error

Complexity

Floating point operation (flop) count

Asymptotic estimate as function of problem size

$$O(n^3) = k_3 n^3 + k_2 n^2 + \cdots$$

notation $O(.)$, order of, indicates dominating term

For example, discrete Fourier transform (DFT) of n points, implemented directly, has complexity $O(n^2)$, whereas fast Fourier transform algorithm, `fft` in Matlab, has complexity $O(n \log n)$

Algorithms `svd`, `eig`, `schur`, `chol` have all complexity $O(n^3)$, but their constant factor k_3 varies

Complexity

With numerical computation, complexity estimates may also involve the required accuracy ϵ, for example $O(\sqrt{n} \log \epsilon^{-1})$

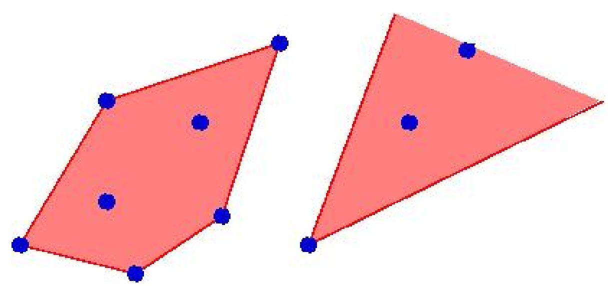
When asymptotic complexity is a polynomial function of problem size, we say that the algorithm is polynomial-time

When no polynomial-time algorithms exist, there may still be exponential-time algorithms to solve the problem, for example finding 0/1 solutions to a linear system of equations with integer data

Convexity and cones

Set K is convex if the line segment between any two points in K lies in K: $\forall x_1, x_2 \in K,\ \lambda x_1 + (1-\lambda)x_2 \in K,\ \forall\lambda,\ 0 \le \lambda \le 1$

The convex hull of a set K is the set of all convex combinations of points in K: $\operatorname{conv} K = \{\sum_i \lambda_i x_i \ :\ x_i \in K,\ \lambda_i \ge 0,\ \sum_i \lambda_i = 1\}$



A set K is a cone if for every $x \in K$ and $\lambda \ge 0$ we have $\lambda x \in K$

So convex cones are invariant under addition and multiplication by a nonnegative constant

Open and closed sets

A set K is open if, when starting from any point in K, one can move by a small amount in any direction while staying in K

A set K is closed if its complement is open

A set $K \subset \mathbb{R}^n$ is called compact if it is closed and bounded

Examples:
- the interval $[0, 1]$ is closed in $\mathbb{R}$
- the interval $(0, 1)$ is open is $\mathbb{R}$
- the interval $[0, 1)$ is neither open nor closed in $\mathbb{R}$
- the empty set is both open and closed (clopen)
- the set of rational numbers between 0 and 1 is closed in $\mathbb{Q}$ but not in $\mathbb{R}$

Dual cone

Let $K \subset \mathbb{R}^n$ be a cone

Then the set

$$K^\star = \{y \in \mathbb{R}^n \ : \ y^T x \geq 0 \text{ for all } x \in K\}$$

is its dual cone

$K^\star$ can be viewed as the set of nonnegative linear maps on K

$K^\star$ is always a closed convex cone

If K itself is a closed convex cone, then $K^{\star\star} = K$

Dual set

Let $K \subset \mathbb{R}^n$ be a set containing the origin

Then the set

$$K^o = \{y \in \mathbb{R}^n \ : \ y^T x \leq 1 \text{ for all } x \in K\}$$

is its polar set, and the set

$$K^\star = -K^o = \{y \in \mathbb{R}^n \ : \ 1 + y^T x \geq 0 \text{ for all } x \in K\}$$

is its dual set

Compare with dual cone: nonhomogeneous coordinates