Advanced Numerical Analysis

Lecture 10
Spring 2025

EPFL

Daniel Kressner

Quiz of Exercise Set 9

(a) For *any* invertible matrix A, right-hand side $\mathbf{b}$, and starting vector $\mathbf{x}_0$, there is a choice of α such that the Richardson method converges.

- True
- False

(b) Consider a family of linear systems

$$A_n\mathbf{x} = \mathbf{b}_n, \quad A_n \in \mathbb{R}^{n\times n},$$

such that

- A_n is symmetric positive definite;
- $\kappa_2(A_n) = \|A_n\|_2\|A_n^{-1}\|_2 = O(n^2)$ for $n \to \infty$;
- $\|\mathbf{x}\|_2 = 1$.

Consider fixed accuracy $\varepsilon > 0$. Let k_n denote the minimal number of iterations of the Richardson method (with optimal α, zero starting vector, no preconditioner) needed to attain $\|\mathbf{x}_{k_n} - \mathbf{x}\|_2 \leq \varepsilon$. Then for $n \to \infty$ it holds that

- $k_n = O(1)$
- $k_n = O(\log n)$
- $k_n = O(n)$
- $k_n = O(n^2)$

Quiz of Exercise Set 9

(c) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable on $\mathbb{R}^n$. If $\mathbf{x}$ is a minimum of f then $\nabla f(\mathbf{x}) = 0$.

- ○ True
- ○ False

(d) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable on $\mathbb{R}^n$ and $\mathbf{x}$ such that $\nabla f(\mathbf{x}) \neq 0$. Then for every $\varepsilon > 0$ there is $\mathbf{y}$ with $\|\mathbf{y} - \mathbf{x}\| \leq \varepsilon$ and $f(\mathbf{y}) < f(\mathbf{x})$.

- ○ True
- ○ False

CG method

Given $\mathbf{x}^{(0)} \in \mathbb{R}^n$, let $\mathbf{r}^{(0)} = \mathbf{b} - A\mathbf{x}^{(0)}$ and $\mathbf{p}^{(0)} = \mathbf{r}^{(0)}$. Then for all $k \geq 0$,

$$\begin{cases} \alpha_k = \frac{\langle \mathbf{p}^{(k)}, \mathbf{r}^{(k)} \rangle}{\langle \mathbf{p}^{(k)}, A\mathbf{p}^{(k)} \rangle}; \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}^{(k)}; \\ \mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A\mathbf{p}^{(k)}; \\ \beta_k = \frac{\langle \mathbf{r}^{(k+1)}, A\mathbf{p}^{(k)} \rangle}{\langle \mathbf{p}^{(k)}, A\mathbf{p}^{(k)} \rangle}; \\ \mathbf{p}^{(k+1)} = \mathbf{r}^{(k+1)} - \beta_k \mathbf{p}^{(k)}. \end{cases}$$

Convergence of CG

Theorem (Theorem 5.9)

Let $A \in \mathbb{R}^{n \times n}$ *be SPD. Then CG yields after at most n iterations the exact solution (assuming exact arithmetic).*

- Usually not very relevant because: (1) One hopes to get good accuracy well before. (2) Result (miserably) fails to hold in floating point arithmetic.
- Exception: Solving sparse linear systems over finite fields [Teitelbaum'1998].

Convergence of CG

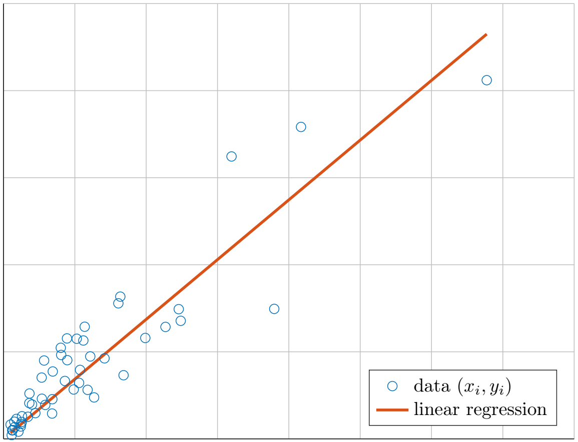
Theorem (Theorem 5.10)

Let $A \in \mathbb{R}^{n\times n}$ be SPD and consider linear system $A\mathbf{x} = \mathbf{b}$. For $k \geq 0$, let $\mathbf{e}^{(k)} := \mathbf{x}^{(k)} - \mathbf{x} \in \mathbb{R}^n$, where $\mathbf{x}^{(k)}$ is the kth iterate of CG. Then,

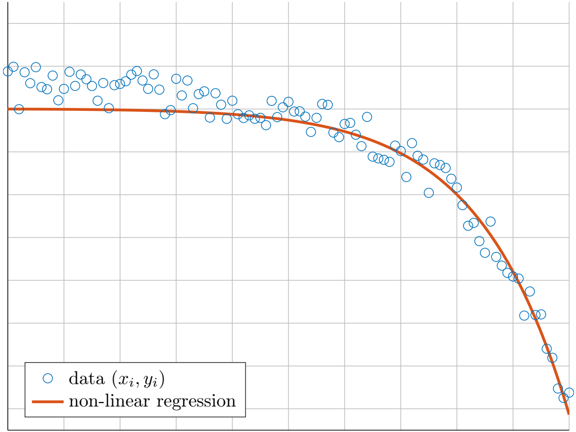
$$\left\|\mathbf{e}^{(k)}\right\|_A \leq 2\frac{C^k}{1+C^{2k}}\left\|\mathbf{e}^{(0)}\right\|_A, \quad \textit{with} \quad C := \frac{\sqrt{\kappa_2(A)}-1}{\sqrt{\kappa_2(A)}+1}.$$

- Reduces $\kappa_2(A)$ (Gradient descent) to $\sqrt{\kappa_2(A)}$ (CG)
- Preconditioning can be used to reduce $\sqrt{\kappa_2(A)}$ further.

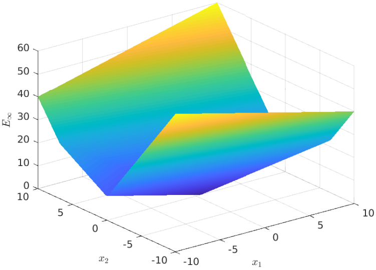
Linear regression

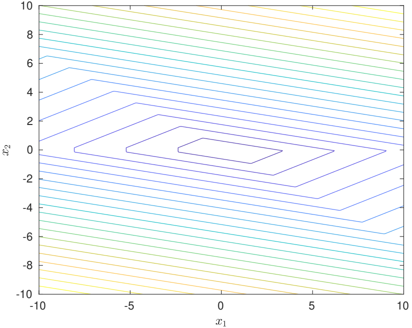


Nonlinear regression

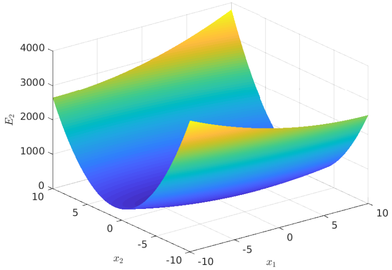


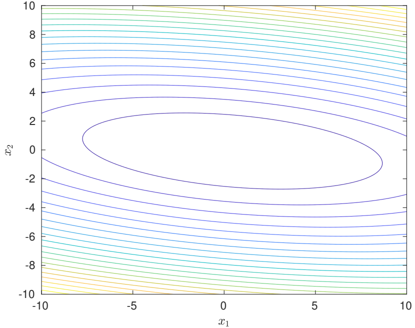
Maximum norm error function





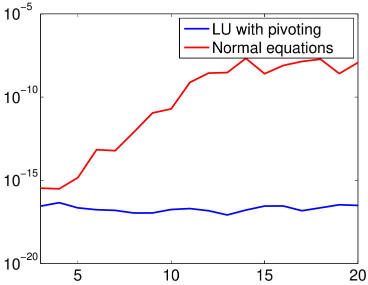
Euclidean norm error function





Numerical instability of normal equations

Setting: $n \times n$ linear system $A\mathbf{x} = \mathbf{b}$. Condition number of Vandermonde matrix A grows exponentially with n!



Residual norm vs. n when solving linear system by LU with pivoting or normal equations by Cholesky.