

Basic linear algebra recap. Convergence rates. Line Search

Seminar

Optimization for ML. Faculty of Computer Science. HSE University

Basic linear algebra recap

- Naive matmul $\mathcal{O}(n^3)$, naive matvec $\mathcal{O}(n^2)$

Basic linear algebra recap

- Naive matmul $\mathcal{O}(n^3)$, naive matvec $\mathcal{O}(n^2)$
- All matrices have SVD

$$A = U\Sigma V^T$$

Basic linear algebra recap

- Naive matmul $\mathcal{O}(n^3)$, naive matvec $\mathcal{O}(n^2)$
- All matrices have SVD

$$A = U\Sigma V^T$$

- $\text{tr}(ABCD) = \text{tr}(DABC) = \text{tr}(CDAB) = \text{tr}(BCDA)$ for any matrices ABCD if the multiplication defined.

Basic linear algebra recap

- Naive matmul $\mathcal{O}(n^3)$, naive matvec $\mathcal{O}(n^2)$
- All matrices have SVD

$$A = U\Sigma V^T$$

- $\text{tr}(ABCD) = \text{tr}(DABC) = \text{tr}(CDAB) = \text{tr}(BCDA)$ for any matrices ABCD if the multiplication defined.
- $\langle A, B \rangle = \text{tr}(A^T B)$

Convergence rate

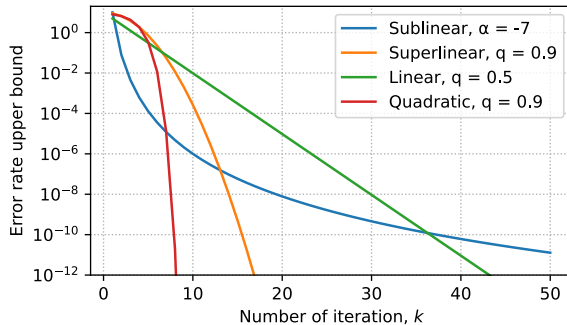


Figure 1: Illustration of different convergence rates

- Linear (geometric, exponential) convergence:

$$r_k \leq Cq^k, \quad 0 < q < 1, C > 0$$

Convergence rate



Figure 1: Illustration of different convergence rates

- Linear (geometric, exponential) convergence:

$$r_k \leq Cq^k, \quad 0 < q < 1, C > 0$$

- Any convergent sequence, that is slower (faster) than any linearly convergent sequence has sublinear (superlinear) convergence

Root test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of non-negative numbers, converging to zero, and let

$$q = \lim_{k \rightarrow \infty} \sup_k r_k^{1/k}$$

- If $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .

Root test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of non-negative numbers, converging to zero, and let

$$q = \lim_{k \rightarrow \infty} \sup_k r_k^{1/k}$$

- If $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.

Root test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of non-negative numbers, converging to zero, and let

$$q = \lim_{k \rightarrow \infty} \sup_k r_k^{1/k}$$

- If $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.
- If $q = 1$, then $\{r_k\}_{k=m}^{\infty}$ has sublinear convergence.

Root test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of non-negative numbers, converging to zero, and let

$$q = \lim_{k \rightarrow \infty} \sup_k r_k^{1/k}$$

- If $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.
- If $q = 1$, then $\{r_k\}_{k=m}^{\infty}$ has sublinear convergence.
- The case $q > 1$ is impossible.

Ratio test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of strictly positive numbers converging to zero. Let

$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

- If there exists q and $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .

Ratio test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of strictly positive numbers converging to zero. Let

$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

- If there exists q and $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.

Ratio test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of strictly positive numbers converging to zero. Let

$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

- If there exists q and $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.
- If q does not exist, but $q = \lim_{k \rightarrow \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with a constant not exceeding q .

Ratio test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of strictly positive numbers converging to zero. Let

$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

- If there exists q and $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.
- If q does not exist, but $q = \lim_{k \rightarrow \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with a constant not exceeding q .
- If $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$, then $\{r_k\}_{k=m}^{\infty}$ has sublinear convergence.

Ratio test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of strictly positive numbers converging to zero. Let

$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

- If there exists q and $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.
- If q does not exist, but $q = \lim_{k \rightarrow \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with a constant not exceeding q .
- If $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$, then $\{r_k\}_{k=m}^{\infty}$ has sublinear convergence.
- The case $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} > 1$ is impossible.

Ratio test

Let $\{r_k\}_{k=m}^{\infty}$ be a sequence of strictly positive numbers converging to zero. Let

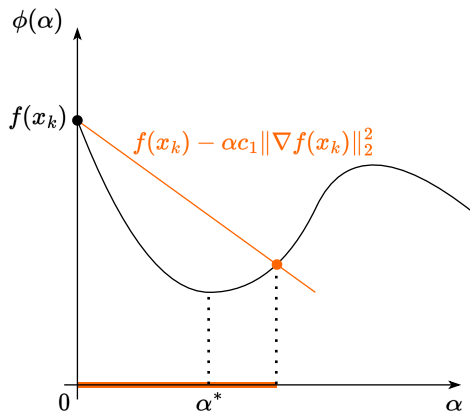
$$q = \lim_{k \rightarrow \infty} \frac{r_{k+1}}{r_k}$$

- If there exists q and $0 \leq q < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with constant q .
- In particular, if $q = 0$, then $\{r_k\}_{k=m}^{\infty}$ has superlinear convergence.
- If q does not exist, but $q = \lim_{k \rightarrow \infty} \sup_k \frac{r_{k+1}}{r_k} < 1$, then $\{r_k\}_{k=m}^{\infty}$ has linear convergence with a constant not exceeding q .
- If $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} = 1$, then $\{r_k\}_{k=m}^{\infty}$ has sublinear convergence.
- The case $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} > 1$ is impossible.
- In all other cases (i.e., when $\lim_{k \rightarrow \infty} \inf_k \frac{r_{k+1}}{r_k} < 1 \leq \lim_{k \rightarrow \infty} \sup_k \frac{r_{k+1}}{r_k}$) we cannot claim anything concrete about the convergence rate $\{r_k\}_{k=m}^{\infty}$.

Line search

Typical line search problem is finding the good value α of the stepsize:

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$



Line search methods

- Solution localization methods:

Line search methods

- Solution localization methods:
 - Dichotomy search method

Line search methods

- Solution localization methods:
 - Dichotomy search method
 - Golden selection search method

Line search methods

- Solution localization methods:
 - Dichotomy search method
 - Golden selection search method
- Inexact line search:

Line search methods

- Solution localization methods:
 - Dichotomy search method
 - Golden selection search method
- Inexact line search:
 - Sufficient decrease

Line search methods

- Solution localization methods:
 - Dichotomy search method
 - Golden selection search method
- Inexact line search:
 - Sufficient decrease
 - Goldstein conditions

Line search methods

- Solution localization methods:
 - Dichotomy search method
 - Golden selection search method
- Inexact line search:
 - Sufficient decrease
 - Goldstein conditions
 - Curvature conditions

Line search methods

- Solution localization methods:
 - Dichotomy search method
 - Golden selection search method
- Inexact line search:
 - Sufficient decrease
 - Goldstein conditions
 - Curvature conditions
 - The idea behind backtracking line search

Problem 1. Stupid important idea on matrix computations.

Suppose, you have the following expression

$$b = A_1 A_2 A_3 x,$$

where the $A_1, A_2, A_3 \in \mathbb{R}^{3 \times 3}$ - random square dense matrices and $x \in \mathbb{R}^n$ - vector. You need to compute b .

Which one way is the best to do it?

1. $A_1 A_2 A_3 x$ (from left to right)

Check the simple code snippet after all.

Problem 1. Stupid important idea on matrix computations.

Suppose, you have the following expression

$$b = A_1 A_2 A_3 x,$$

where the $A_1, A_2, A_3 \in \mathbb{R}^{3 \times 3}$ - random square dense matrices and $x \in \mathbb{R}^n$ - vector. You need to compute b .

Which one way is the best to do it?

1. $A_1 A_2 A_3 x$ (from left to right)
2. $(A_1 (A_2 (A_3 x)))$ (from right to left)

Check the simple  code snippet after all.

Problem 1. Stupid important idea on matrix computations.

Suppose, you have the following expression

$$b = A_1 A_2 A_3 x,$$

where the $A_1, A_2, A_3 \in \mathbb{R}^{3 \times 3}$ - random square dense matrices and $x \in \mathbb{R}^n$ - vector. You need to compute b .

Which one way is the best to do it?

1. $A_1 A_2 A_3 x$ (from left to right)
2. $(A_1 (A_2 (A_3 x)))$ (from right to left)
3. It does not matter

Check the simple  code snippet after all.

Problem 1. Stupid important idea on matrix computations.

Suppose, you have the following expression

$$b = A_1 A_2 A_3 x,$$

where the $A_1, A_2, A_3 \in \mathbb{R}^{3 \times 3}$ - random square dense matrices and $x \in \mathbb{R}^n$ - vector. You need to compute b .

Which one way is the best to do it?

1. $A_1 A_2 A_3 x$ (from left to right)
2. $(A_1 (A_2 (A_3 x)))$ (from right to left)
3. It does not matter
4. The results of the first two options will not be the same.

Check the simple 📄code snippet after all.

Problem 2. Connection between Frobenius norm and singular values.

Let $A \in \mathbb{R}^{m \times n}$, and let $q := \min\{m, n\}$. Show that

$$\|A\|_F^2 = \sum_{i=1}^q \sigma_i^2(A),$$

where $\sigma_1(A) \geq \dots \geq \sigma_q(A) \geq 0$ are the singular values of matrix A . Hint: use the connection between Frobenius norm and scalar product and SVD.

Problem 3. Known your inner product.

Simplify the following expression:

$$\sum_{i=1}^n \langle S^{-1} a_i, a_i \rangle,$$

where $S = \sum_{i=1}^n a_i a_i^T$, $a_i \in \mathbb{R}^n$, $\det(S) \neq 0$

Problem 4. Simple convergence rates

Determine the convergence or divergence of the given sequences:

- $r_k = \frac{1}{3^k}$

Problem 4. Simple convergence rates

Determine the convergence or divergence of the given sequences:

- $r_k = \frac{1}{3^k}$
- $r_k = \frac{4}{3^k}$

Problem 4. Simple convergence rates

Determine the convergence or divergence of the given sequences:

- $r_k = \frac{1}{3^k}$
- $r_k = \frac{4}{3^k}$
- $r_k = \frac{1}{k^{10}}$

Problem 4. Simple convergence rates

Determine the convergence or divergence of the given sequences:

- $r_k = \frac{1}{3^k}$
- $r_k = \frac{4}{3^k}$
- $r_k = \frac{1}{k^{10}}$
- $r_k = 0.707^k$

Problem 4. Simple convergence rates

Determine the convergence or divergence of the given sequences:

- $r_k = \frac{1}{3^k}$
- $r_k = \frac{4}{3^k}$
- $r_k = \frac{1}{k^{10}}$
- $r_k = 0.707^k$
- $r_k = 0.707^{2^k}$

Problem 5. One test is simpler, than another.

Determine the convergence or divergence of the following sequence:

$$r_k = \frac{1}{k^k}$$

Problem 6. Quadratic convergence.

Show, that the following sequence does not have a quadratic convergence.

$$r_k = \frac{1}{3^{k^2}}$$