6. Matrix operations

Let K be a field and assume $+, -, \times, ^{-1}$, $\mathbb{Z} \to K$ can be done in $O(1)$.

6.1. Multiplication

Q How quickly can we multiply two $n \times n$-matrices?

Rmk Triv. alg.: $O(n^3)$

Thm 6.1.1 (Strassen, Gaussian Elimination is not optimal)
You can multiply $A, B \in M_{n\times n}(K)$ in time $O(n^{\log_2 7})$ (on an $O(\log n)$-bit RAM).

Idea w.l.o.g. $n = 2^k$.

Write $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$,

with $\frac{n}{2} \times \frac{n}{2}$-matrices A_{ij}, B_{ij}.

$$\Rightarrow AB = \underbrace{\begin{bmatrix} \sum_{j=1}^{2} A_{1j} B_{j1} & \sum A_{1j} B_{j2} \\ \sum A_{2j} B_{j1} & \sum A_{2j} B_{j2} \end{bmatrix}}_{\substack{\text{total: 8 mult. of } \frac{n}{2}\times\frac{n}{2}\text{-matrices} \\ \rightsquigarrow \text{ time } O(n^{\log_2 8}) = O(n^3).}}.$$

Actually, 7 mult. are enough! [Similar to Karatsuba!]

and some number of additions/subtractions

$$\Rightarrow \text{total time} \approx 2^{2k} + 7 \cdot 2^{2(k-1)} + 7^2 \cdot 2^{2(k-2)} + \dots + 7^k \approx 7^k = n^{\log_2 7}.$$

$\square$

<u>Rmks</u> The exponent ω s.t. time $O(n^{\omega})$ suffices has been improved many times.

Strassen: $\omega = \log_2 7 \approx 2.807$

Current record: $\omega \lesssim 2.373$

It's very unclear if $O_{\varepsilon}(n^{2+\varepsilon})$ is possible for all $\varepsilon > 0$.

6.2. Determinant, rank, inverse

(Strassen, Gaussian Elimination is not Optimal)

Thm 6.2.1 Assume we can multiply $n \times n$-matrices in $O(n^{\omega})$, with $\omega > 2$.

Then, given an $n \times n$-matrix A, we can compute an invertible $n \times n$-matrix B and its determinant (its inverse B^{-1}) such that BA is in reduced row echelon form in time $O_{\omega}(n^{\omega})$.

first nonzero entry in each row is 1. Entries above and below this 1 are 0. Each row has at least as many leading zeros as the previous row.

Rmk Gaussian elimination does this in $O(n^2 m)$ for an $n \times m$-matrix A (and $n \times n$-matrix B).

Cor 6.2.2 We can compute $\det(A)$, $\mathrm{rk}(A)$, A^{-1}, bases for $\ker(A)$, $\mathrm{im}(A)$, in $O_{\omega}(n^{\omega})$.

Pf of Cor $\det(A) = \det(B)^{-1} \cdot \det(AB) = \det(B)^{-1} \cdot$ prod. of diagonal entries of AB

$\mathrm{rk}(A) =$ number of nonzero rows in A

$\boxed{AB \text{ is upper triangular}}$

If $\mathrm{rk}(A) = n$, then $AB = I_n$, so $A^{-1} = B$.

$\ker(A) = \ker(BA) = \ldots$

$\mathrm{im}(A) = B^{-1} \cdot \mathrm{im}(BA) = \ldots$

$\square$

Alg for Thm W.l.o.g $n=2^k$.

1) Find $B_1=\begin{bmatrix}* & 0\\ 0 & I\end{bmatrix}$ s.t. $B_1A=\begin{bmatrix}\text{RREF} & *\\ * & *\end{bmatrix}$

by recursively applying the alg. to the top left $\frac{n}{2}\times\frac{n}{2}$-matrix

2) Find $B_2=\begin{bmatrix}I & 0\\ * & I\end{bmatrix}$ s.t. in B_2B_1A ~~above and~~ below any leading 1 in the top left, there are just 0s in the bottom left.

3) Find $B_3=\begin{bmatrix}I & 0\\ 0 & *\end{bmatrix}$ s.t. $B_3B_2B_1A=\begin{bmatrix}\text{RREF} & *\\ \text{RREF} & *\end{bmatrix}$

by recursively applying the alg. to the bottom left $\frac{n}{2}\times\frac{n}{2}$-matrix.

4) Find $B_4=\begin{bmatrix}I & *\\ 0 & I\end{bmatrix}$ s.t. in $B_4\cdots B_1A$ above any leading 1s in the bottom left, there are just 0s in the top left.

~~5) Find a permutation matrix B_5 s.t. in $B_5\cdots B_1A$, the number of leading 0s among the first $\frac{n}{2}$ columns is non decreasing. Then, the left half of $B_5\cdots B_1A$ is in RREF~~

5) Apply steps 1–4 to the right half of the matrix, ignoring all rows that have nonzero entries in the left half.

6) Apply a permutation matrix to ensure that the number of leading 0s in each row is non-decreasing as you move downwards.

The resuling matrix is in RREF.

Total: 4 recursive calls with $\frac{n}{2}\times\frac{n}{2}$-matrices and a bdd. nr. of mult. of $n\times n$-matrices

$\Rightarrow$ Time $\lesssim 2^{k\omega}+4\cdot 2^{(k-1)\omega}+\dots+4^k \lesssim_\omega 2^{k\omega}=n^\omega$. $\square$

6.3. Characteristic polynomial

Thm 6.3.1 (Hessenberg, cf. section 2.2.4 in Cohen)

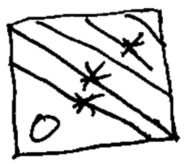
We can compute the char. pol. $\chi_A(x) = \det(xI_n - A)$ of an $n\times n$-matrix A in $O(n^3)$.

Rmk This can also be done in $O_\omega(n^\omega \log n)$ (Keller-Gehrig, Fast algorithms for the char. pol.) and randomizedly in $O_\omega(n^\omega)$ (Pernet-Storjohann, Faster alg. for the char. pol.) if $\#K \geq 2n^2$.

Pf First, find a similar matrix $B = (b_{ij})_{i,j}$ in

Hessenberg form: $b_{ij} = 0 \quad \forall 1\leq i,j\leq n$ such that $i \geq j+2$

(Lemma 6.3.2)



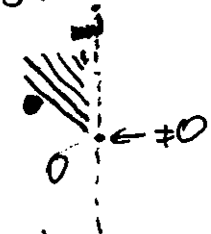
Then, $\chi_A(x) = \chi_B(x)$ ~~$\det(xI_n - A) = \det(xI_n - B)$~~ can be computed in $O(n^3)$ using Laplace expansion. (Lemma 6.3.3)

$\square$

Lemma 6.3.2 You can compute a matrix B in Hessenberg form which is similar to $A = (a_{ij})_{i,j}$ in $O(n^3)$.

Alg Start with $B = A$. We'll "fix" the columns starting from the left.

For $j = 1, \dots, n-1$:



If $b_{ij} \neq 0$ for some $i \geq j+2$:

Let i_0 be the smallest $i \geq j+1$ s.t. $b_{ij} \neq 0$.

Exchange rows $j+1$ and i_0

and columns $j+1 (>j)$ and $i_0 (>j)$.

(Then, $b_{j+1,j} \neq 0$.)

For each $i \geq j+2$:

Subtract $v := \frac{b_{ij}}{b_{j+1,j}}$ times row $j+1$ from row i and

add v times column $i^{(>j)}$ to column $j+1 (>j)$.

(Then, $b_{i,j} = 0$.)

□

Lemma 6.3.3

~~You can compute $\det(xI_n - B)$ in $O(n^3)$.~~

Pf ~~Let~~ Let $B=(b_{ij})$ be an $n\times n$-matrix in Hessenberg form and let B_m be its top left $m\times m$ minor for $0\le m\le n$.
Write $p_m(x)=\chi_{B_m}(x)$.

~~Then, $\chi_B(x) = (x-b_{nn})\chi_{B_{m-1}}$~~

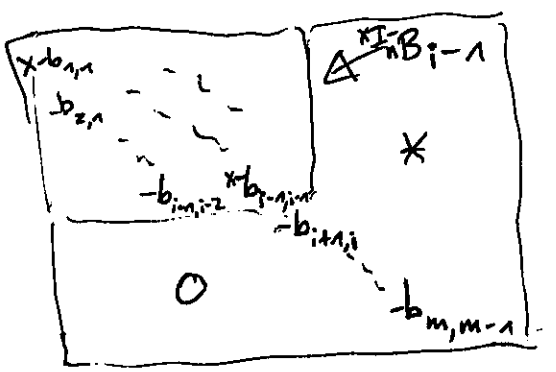
Then, $$p_m(x)=(x-b_{mm})p_{m-1}(x)-\sum_{i=1}^{m-1} b_{im}\cdot(b_{i+1,i}\cdots b_{m,m-1})\cdot p_{i-1}(x).$$

Rmk Since $\deg(p_m)=m$, we can compute $p_0,\dots,p_n$ in $O(n^3)$.

Pf Use Laplace expansion on column m. This involves picking a row.

Row m gives $(x-b_{mm})p_{m-1}(x)$.

Removing row $1\le i<m$ and column m leaves the matrix



which produces the summand

$$(-1)^{i+m}\cdot(-b_{im})(-b_{i+1,i})\cdots(-b_{m,m-1})\cdot p_{i-1}(x)$$
$$= -\,b_{im}\,b_{i+1,i}\cdots b_{m,m-1}\cdot p_{i-1}(x).$$

$\square$