Chapter 12 Fast Linear Algebra

12.1 Strassen's matrix multiplication

Algorithm 12.1 Matrix multiplication Algorithm

Input: $A, B \in R^{n \times n}$ where R is a ring and $n = 2^k$ for some $k \in \mathbb{N}$

Output: The product matrix $A \cdot B \in R^{n \times n}$.

1. if $n = 1$ then let $A = (a), B = (b)$ for some $a, b \in R$ and return (ab)
2. write $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}, \end{pmatrix}, B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22}, \end{pmatrix}$

3.

$$
\begin{array}{lclclcl}
S_1 & \longleftarrow & A_{21} + A_{22}, & T_1 & \longleftarrow & B_{12} - B_{11} \\
S_2 & \longleftarrow & S_1 - A_{11}, & T_2 & \longleftarrow & B_{22} - T_1 \\
S_3 & \longleftarrow & A_{11} - A_{21}, & T_3 & \longleftarrow & B_{22} - B_{12} \\
S_4 & \longleftarrow & A_{12} - S_2, & T_4 & \longleftarrow & T_2 - B_{21}
\end{array}
$$

4. call the algorithm recursively to compute

$$
\begin{array}{lclclcl}
P_1 & = & A_{11}B_{11}, & P_5 & = & S_1T_1, \\
P_2 & = & A_{12}B_{21}, & P_6 & = & S_2T_2, \\
P_3 & = & S_4B_{22}, & P_7 & = & S_3T_3, \\
P_4 & = & A_{22}T_4, & & &
\end{array}
$$

5.

$$\begin{array}{rclrcl} U_1 & \longleftarrow & P_1 + P_2, & U_5 & \longleftarrow & U_4 + P_3 \\ U_2 & \longleftarrow & P_1 + P_6, & U_6 & \longleftarrow & U_3 - P_4 \\ U_3 & \longleftarrow & U_2 + P_7, & U_7 & \longleftarrow & U_3 + P_5 \\ U_4 & \longleftarrow & U_2 + P_5 & & & \end{array}$$

6. return $\begin{pmatrix} U_1 & U_5 \\ U_6 & U_7 \end{pmatrix}$

Theorem 1. *Algorithm 12.1 correctly computes the product matrix and uses at most $6n^{\log_2 7}$ additions and multiplications in R. For arbitrary $n \in \mathbb{N}$, an $n \times n$ matrix product can be computed with $42n^{\log_2 7}$ or $O(n^{\log_2 7})$ ring operations.*

12.2 Applications: fast modular composition of polynomials

12.3 Linearly recurrent sequences

Black box linear algebra: A square matrix $A \in F^{n \times n}$ is not given by its $n \times n$ array of entries, but by a black box for evaluating A, that is, a procedure which on input of a vector $v \in F^n$ produces $Av \in F^n$.

We denote by $c(A)$ the arithmetic cost for evaluting A at a vector Av.

class of matrices	$c(A)$
general	$2n^2 - n$
Sylvester matrix	$O(M(n))$
DFT_ω, ω a primitive 2^kth root of unity	$O(n \log n)$
Vandermonde matrix	$O(M(n) \log n)$
Berlekamp matrix over F_q	$O(M(n) \log q)$
sparse, with $\leq s$ nonzero entries	$\leq 2s$
random squares integer factorization	$O(n \log^2 n)$

Definition 1. *A sequence $a = (a_i)_{i\in\mathbb{N}} \in V^{\mathbb{N}}$ is linearly recurrent if there exist $n \in \mathbb{N}$ and $f_0, \ldots, f_n \in F$ with $f_n \neq 0$ such that*

$$\sum_{0\leq j\leq n} f_j a_{i+j} = f_n a_{i+n} + \cdots + f_1 a_{i+1} + f_0 a_i = 0$$

for all $i \in \mathbb{N}$. The polynomial $f = \sum_{0\leq j\leq n} f_j x^j \in F[x]$ of degree n is called a characteristic(or generating) polynomial of a.

1. $V = F = \mathbb{Q}$, $a_0 = 0, a_1 = 1, a_{i+2} = a_{i+1} + a_i$ for all $i \geq 2$. Then $a = (a_i)_{i \geq 0}$ is the Fibonacci sequence, and $f = x^2 - x - 1$ is a characteristic polynomial of a

2. $V = F^n, A \in F^{n \times n}$ and $b \in F^n$ arbitrary. Then $a = (A^i b)_{i \in \mathbb{N}}$ is linearly recurrent, and any characteristic polynomial of $(A^i)_{i \in \mathbb{N}}$ also annihilates a. The subspace of F^n spanned by the entries of a is called the Krylov subspace of A and b

3. $V = F, A \in F^{n \times n}$ and $b, u \in F^n$ arbitrary. Then $a = (u^T A^i b)_{i \in \mathbb{N}}$ is linearly recurrent, and any characteristic polynomial of $(A^i)_{i \in \mathbb{N}}$ also annihilates a.

For $f \in \sum_{0\le i\le n} f_j x^j \in F[x]$ and $a = (a_i)_{i\in\mathbb{N}} \in V^{\mathbb{N}}$, we set

$$f \bullet a = \left(\sum_{0\le i\le n} f_j a_{i+j} \right)_{i\in\mathbb{N}} \in V^{\mathbb{N}}.$$

The set of all characteristic polynomial of $a \in V^{\mathbb{N}}$ together with the zero polynomial is an ideal in F[x]:

$$\mathsf{Ann}(a)(annihilator) =< m_a >$$

m_a is called the minimal polynomial of a. The degree of m_a is called the recursion order of a.

$$\begin{aligned} f &= 0 \,\text{or}\, f \,\text{is a characteristic polynomials of}\, a \\ &\iff f \bullet a = 0 \iff f \in \mathsf{Ann}(a) \iff m_a | f, \end{aligned}$$

Lemma 1. *Let $a = (a_i)_{i \in \mathbb{N}} \in F^{\mathbb{N}}$ be linearly recurrent, $h = \sum_{i \in \mathbb{N}} a_i x^i \in F[[x]]$, $f \in F[x]/\ \{0\}$ of degree d, and $r = rev(f)$ is reversal.*

1. *The following are equivalent*
   - *f is a characteristic polynomial of a,*
   - *$r \cdot h$ is a polynomial of degree less than d*
   - *$h = g/r$ for some $g \in F[x]$ with $\deg g < d$.*
2. *If f is the minimal polynomial of a, then $d = \max\{1 + \deg g, \deg r\}$ and $\gcd(g, r) = 1$.*

We can compute m_a by solving the Padé approximate problem

$$h \equiv \frac{s}{t} \mod x^{2n}, x \nmid t, \deg x < n, \deg t \leq n, \gcd(s,t) = 1(*)$$

Algorithm 12.9 Minimal polynomial for $F^{\mathbb{N}}$

Input: An upper bound $n \in \mathbb{N}$ on the recursion order and the first $2n$ entries $a_0, \ldots, a_{2n-1} \in F$ of a linearly recurrent sequence $a \in F^{\mathbb{N}}$.

Output: The minimal polynomial $m_a \in F[x]$ of a.

1. $h \longleftarrow a_{2n-1}x^{2n-1} + \cdots + a_1 x + a_0$, call EEA to compute s, t satisfy (*)
2. $d \longleftarrow \max\{1 + \deg g, \deg r\}$, return rev$_d(t)$.

12.4 Wiedemann's algorithm and black box linear algebra

Question: For a given matrix $A \in F^{n \times m}$ and a vector $b \in F^n$ over a field F, to find one solution $y \in F^m$ to the equation $Ay = b$.

Suppose $m = m_a = \sum_{0 \le j \le d} m_j x^j \in F[x]$ is the minimal polynomial of the linearly recurrent sequence $a = (A^i b)_{i \in \mathbb{N}}$. Then we have

$$m(A)b = \sum_{0 \le j \le d} m_j A^j b = 0 \in F^n$$

If A is nonsingular, then $m_0 = m(0)$ is nonzero,

$$A \cdot (-m_0^{-1}) \sum_{1 \le j \le d} m_j A^{j-1} b = -m_0^{-1} \sum_{1 \le j \le d} m_j A^j b = b,$$

and

$$y = (-m_0^{-1}) \sum_{1 \le j \le d} m_j A^{j-1} b \in F^n$$

is the required solution.

Solving a nonsingular square linear system

Input: A nonsingular matrix $A \in F^{n \times n}$ and a vector $b \in F^n$.

Output: $y = A^{-1}b \in F^n$

1. compute the minimal polynomial $m \in F[x]$ of the linearly recurrent sequence $(A^i b)_{i \in \mathbb{N}} \in (F^n)^{\mathbb{N}}$

2. $h \longleftarrow -\frac{m - m(0)}{m(0) \cdot x} \in F[x]$, compute $y = h(A)b$ in a Horner-like fashion

3. return y

Minimal polynomial for Krylov subspaces

Input: A matrix $A \in F^{n \times n}$ and a vector $b \in F^n$

Output: The minimal polynomial $m \in F[x]$ of the sequence $(A^i b)_{i \in \mathbb{N}}$

1. if $b = 0$ then return 1
2. let $U \subseteq F$ be finite
3. choose $u \in U^n$ uniformly at random and compute $u^T A^i b \in F$ for $0 \leq i < 2n$
4. call algorithm 12.9 to compute the minimal polynomial $m \in F[x]$ of $(u^T A^i b)_{i \in \mathbb{N}} \in F^{\mathbb{N}}$ with recursion bound n
5. if $m(A)b = 0$ then return m else go to 3.

Lemma 2. *Let $U \subset F$ be finite, $A \in F^{n\times n}, b \in F^n$ $\{0\}$., f be the minimal polynomial of $(A^i b)_{i\in\mathbb{N}}$ and $d = \deg f$. Then the probability p that f is the minimal polynomial of the sequence $(u^T A^i b)_{i\in\mathbb{N}}$ for a $u \in U^n$ chosen uniformly at random satisfies $p \geq 1 - d/\#U$.*

Theorem 2. *The expected cost of Algorithm 12.12 and 12.13 is at most $4nc(A) + O(n^2)$ field operations if F has at least $2n$ elements.*