Duality via Farkas' lemma

Theorem (Second variant of Farkas' lemma)

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The system $Ax \leq b$ has a solution if and only if for all $\lambda \geqslant 0$ with $\lambda^T A = 0$ one has $\lambda^T b \geqslant 0$.

Duality via Farkas' lemma

Algorithms and running time analysis

Consider the following algorithm to compute the product of two $n \times n$ matrices $A, B \in \mathbb{Q}^{n \times n}$:

for $i = 1, \ldots, n$
    for $j = 1, \ldots, n$
        $c_{ij} := 0$
        for $k = 1, \ldots, n$
            $c_{ij} := c_{ij} + a_{ik} \cdot a_{kj}$

O-notation

Definition

Let $T, f : \mathbb{N} \to \mathbb{R}_{\geqslant 0}$ be functions. We say

- $T(n) = O(f(n))$, if there exist positive constants $n_o \in \mathbb{N}$ and $c \in \mathbb{R}_{>0}$ with

$$T(n) \leq c \cdot f(n) \text{ for all } n \geqslant n_0.$$

- $T(n) = \Omega(f(n))$, if there exist constants $n_o \in \mathbb{N}$ and $c \in \mathbb{R}_{>0}$ with

$$T(n) \geqslant c \cdot f(n) \text{ for all } n \geqslant n_0.$$

- $T(n) = \Theta(f(n))$ if

$$T(n) = O(f(n)) \text{ and } T(n) = \Omega(f(n)).$$

Example

Example

- $T(n) = 2n^2 + 3n + 1 = O(n^2)$
- $T(n) = \Omega(n^2)$
- $T(n) = \Theta(n^2)$.
- $n^2 \log n = O(n^{2+\epsilon})$ for each $\epsilon > 0$

Efficient algorithm, first definition

Definition (Polynomial-time algorithm)

An algorithm runs in polynomial time, if there exists a constant k such that the algorithm runs in time

$$O(n^k)$$

where n is the length of the input (total number of bits).

Why definition is problematic

Size

Definition

The size of an integer x is $\mathrm{size}(x) = \lceil \log(|x| + 1) \rceil$ and for $x \in \mathbb{Q}$, $\mathrm{size}(x) = \mathrm{size}(p) + \mathrm{size}(q)$, where $x = p/q$ with $p, q \in \mathbb{Z}$, $q \geqslant 1$ and $\gcd(p, q) = 1$.

Polynomial time algorithm

Definition

An algorithm is polynomial time, if there exists a constant k such that the algorithm performs $O(n^k)$ operations on rational numbers whose size is bounded by $O(n^k)$. Here n is the number of bits that encode the input of the algorithm. We say that the algorithm runs in time $O(n^k)$.

Example: Euclidean algorithm

Input: Integers $a \geqslant b \geqslant 0$ not both equal to zero
Output: The greatest common divisor $\gcd(a, b)$

if $(b = 0)$ **return** a
else
    Compute $q, r \in \mathbb{N}$ with $b > r \geqslant 0$ and $a = q \cdot b + r$
                                (division with remainder)
    return $\gcd(b, r)$

Analysis

Determinant

Input: $A \in \mathbb{Q}^{n \times n}$
Output: $\det(A)$
if $(n = 1)$
  return a_{11}
else
  $d := 0$
  for $j = 1, \ldots, n$
    $d := (-1)^{1+j} \cdot \det(A_{1j}) + d$
  return d

Analysis

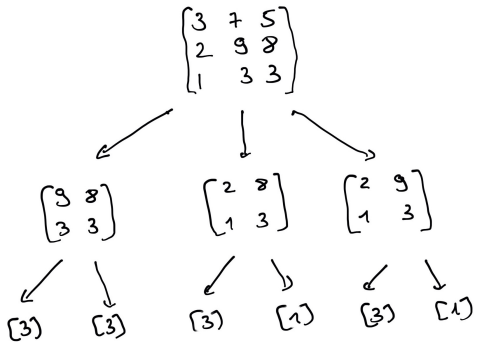


Figure: An example of the recursion tree of the algorithm from Example **??**. The tree corresponds to the run of the algorithm on input $\left(\begin{smallmatrix}3&7&5\\2&9&8\\1&3&3\end{smallmatrix}\right)$.

Gaussian elimination

Input: $A \in \mathbb{Q}^{m \times n}$
Output: A' in row echelon form such that there exists an invertible $Q \in \mathbb{Q}^{m \times m}$ such that $Q \cdot A = A'$.

$A' := A$
$i := 1$
while $(i \leq m)$
    find minimal $1 \leq j \leq n$ such that there exists $k \geqslant i$ such that $a'_{kj} \neq 0$
    If no such element exists, then **stop**
    swap rows i and k in A'
    for $k = i + 1, \ldots, m$
        subtract (a'_{kj}/a'_{ij}) times row i from row k in A'
    $i := i + 1$

Analysis

Theorem

The Gaussian algorithm runs in polynomial time on input $A \in \mathbb{Z}^{m \times n}$. More precisely, the rational numbers produced in the algorithm can be maintained to be ratios of sub-determinants of A' and are thus of polynomial binary encoding length.

Matrix multiplication

If we split the matrices A and B into 4 $n/2 \times n/2$ matrices

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \tag{4}$$

Then

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} A_{11} \cdot B_{11} + A_{12} \cdot B_{21} & A_{11} \cdot B_{12} + A_{12} \cdot B_{22} \\ A_{21} \cdot B_{11} + A_{22} \cdot B_{21} & A_{21} \cdot B_{12} + A_{22} \cdot B_{22} \end{pmatrix}.$$

Strassen's algorithm

$$
\begin{aligned}
M_1 &= (A_{11} + A_{22}) \cdot (B_{11} + B_{22}) \\
M_2 &= (A_{21} + A_{22}) \cdot B_{11} \\
M_3 &= A_{11} \cdot (B_{12} - B_{22}) \\
M_4 &= A_{22} \cdot (B_{21} - B_{11}) \\
M_5 &= (A_{11} + A_{12}) \cdot B_{22} \\
M_6 &= (A_{21} - A_{11}) \cdot (B_{11} + B_{12}) \\
M_7 &= (A_{12} - A_{22}) \cdot (B_{21} + B_{22})
\end{aligned}
$$

.

$$
\begin{aligned}
C_{11} &= M_1 + M_4 - M_5 + M_7 \\
C_{12} &= M_3 + M_5 \\
C_{21} &= M_2 + M_4 \\
C_{22} &= M_1 - M_2 + M_3 + M_6.
\end{aligned}
$$

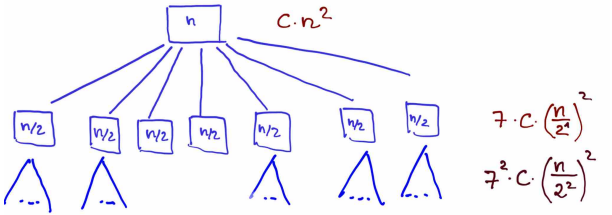
Strassen's algorithm

Input: Two $n \times n$ matrices A and B
Output: $C = \mathrm{FMM}(A, B)$, the product $A \cdot B$

if $n = 1$ return $a_{11} \cdot b_{11}$
else

$M_1 = \mathrm{FMM}(A_{11} + A_{22}, B_{11} + B_{22})$
$M_2 = \mathrm{FMM}(A_{21} + A_{22}, B_{11})$
$M_3 = \mathrm{FMM}A_{11}, B_{12} - B_{22})$
$M_4 = \mathrm{FMM}(A_{22}, (B_{21} - B_{22})$
$M_5 = \mathrm{FMM}(A_{11} + A_{12}, B_{22})$
$M_6 = \mathrm{FMM}(A_{21} - A_{11}, B_{11} + B_{12})$
$M_7 = \mathrm{FMM}(A_{12} - A_{22}, B_{21} + B_{22})$
Compute the matrices $C_{11}, C_{12}, C_{21}, C_{22}$ from $M_1, \ldots, M_7$
return C

Analysis



$$\text{Total: } c \sum_{i=0}^{\log_2 n} 7^i \cdot \frac{n^2}{4^i} = c \cdot n^2 \cdot \sum_{i=0}^{\log_2 n} \left(\frac{7}{4}\right)^i$$

$$\leq c \cdot n^2 \cdot \left(\frac{7}{4}\right)^{\log_2 n} = c \cdot n^{2 + \log_2 \left(\frac{7}{4}\right)} = c \cdot n^{2{,}807}$$

Running time

Theorem (Strassen)

Two $n \times n$ matrices can be multiplied in time (number of arithmetic operations) $O(n^{2+\log_2(7/4)})$.