Lecture 3

Easy cases for computing the permanent

- Borchardt's theorem and its generalization.
- Finding fast algorithms
- Matrices of bounded rank
- Structured matrices

Early efforts on computing the permanent.

H. Minc, *Permanents*, Encyclopedia of Mathematics and its Applications, vol. 6, Addison Wesley 1978, pp. 1-8.

The equation labels refer to these pages

Let $A = (a_{i,j})$ be an $(m \times n)$-matrix, $m \leq n$ over a **commutative ring** $\mathcal{R}$.

Let $\mathcal{S}_{m,n}$ be the set of **injective** functions

$$\sigma : \{1, \ldots, m\} \to \{1, \ldots, n\}$$

and

$$Per(A) = \sum_{\sigma \in \mathcal{S}_{m,n}} \prod_{i=1}^{n} a_{i,\sigma(i)} \qquad (1.1)$$

For the case of $m = n$ we write $per(A)$ instead of $Per(A)$.

Exercise: Compute a few permanents.

Computing $per(A) \cdot det(A)$, I

$$per(A) \cdot det(A) =$$

$$\left(\sum_{\sigma} \underbrace{\prod_{i} a_{i,\sigma(i)}}_{b_\sigma} \right) \cdot \left(\sum_{\sigma} (-1)^{sign(\sigma)} \underbrace{\prod_{i} a_{i,\sigma(i)}}_{b_\sigma} \right) =$$

$$\left(\sum_{\sigma \in Even} b_\sigma + \sum_{\sigma \in Odd} b_\sigma \right) \cdot \left(\sum_{\sigma \in Even} b_\sigma - \sum_{\sigma \in Odd} b_\sigma \right) =$$

$$= \left(\sum_{\sigma \in Even} b_\sigma \right)^2 - \left(\sum_{\sigma \in Odd} b_\sigma \right)^2 \qquad (3.1)$$

Computing $per(A) \cdot det(A)$, II

For square matrices A, B we define
the *Hadamard product* $C = A \circ B$
by $c_{i,j} = a_{i,j} \cdot b_{i,j}$.

Similarly, $A \circ A = A^{(2)}$, and
$B = A^{(k)}$ is defined by $b_{i,j} = a_{i,j}^k$ for $k \in \mathbb{Z}$.

We now write

$$per(A) \cdot det(A) =$$

$$\underbrace{\sum_{\sigma \in Even} \prod_{i} a_{i,\sigma(i)}^2 - \sum_{\sigma \in Odd} \prod_{i} a_{i,\sigma(i)}^2}_{det(A^{(2)})} + f(A) \qquad (3.2)$$

What can we say about $f(A)$?

Computing $per(A) \cdot det(A)$, III

We compute $f(A)$ for $(n \times n)$-matrices,

$n = 1, 2, 3, \dots$

$n = 1$: Trivially,

$$per(A) \cdot det(A) = a_{1,1}^2 = det(A^{(2)})$$

$n = 2$: Also

$$per(A) \cdot det(A) = det(A^{(2)})$$

because

$$per(A) = a_{1,1}a_{2,2} + a_{1,2}a_{2,1}$$

$$det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$$

and

$$per(A) \cdot det(A) = (a_{1,1}a_{2,2})^2 - (a_{1,2}a_{2,1})^2 = $$
$$= a_{1,1}^2 a_{2,2}^2 - a_{1,2}^2 a_{2,1}^2 = det(A^{(2)})$$

Computing $per(A) \cdot det(A)$, IV

For $n = 1, 2$ we have that $f(A) \equiv 0$.

Theorem:(Cayley, 1859)
For $n = 3$ we have

$$per(A) \cdot det(A) = det(A^{(2)}) + f(A)$$

with

$$f(A) = 2 \left(\prod_{i,j} a_{i,j} \right) \cdot det(A^{(-1)})$$

Corollary:
If $n = 3$, $A^{(-1)}$ well defined and $det(A^{(-1)}) = 0$ then $f(A) \equiv 0$ and again

$$per(A) \cdot det(A) = det(A^{(2)})$$

Computing $per(A) \cdot det(A)$, V

In general $f(A)$ is given by

$$f(A) = \left(\sum_{\sigma \neq \tau \in Even} b_\sigma \cdot b_\tau - \sum_{\sigma \neq \tau \in odd} b_\sigma \cdot b_\tau \right) =$$
$$\sum_{\sigma \neq \tau \in Even} \left(\prod_i a_{i,\sigma(i)} a_{i,\tau(i)} \right) - \sum_{\sigma \neq \tau \in Odd} \left(\prod_i a_{i,\sigma(i)} a_{i,\tau(i)} \right)$$

Hence $f(A)$ is a polynomial of degree $2n$.

Cayley's

$$f(A) = 2 \left(\prod_{i,j} a_{i,j} \right) \cdot det(A^{(-1)})$$

is a polynomial of degree $n^2 - n$.

But $n^2 - n = 2n$ implies that $n = 3$, hence Cayley's formula for $f(A)$ does not generalize.

Computing $per(A) \cdot det(A)$, VI

Theorem:(Borchardt, 1855)

Let $\bar{x}, \bar{y}$ be two n-vectors and $A = A(\bar{x}, \bar{y})$ is the $(n \times n)$-matrix with $a_{i,j} = (x_i - y_j)^{-1}$. Then

$$per(A) \cdot det(A) = det(A^{(2)})$$

or, again $f(A) \equiv 0$.

Such matrices are also called *Cauchy matrices*.

- The only $(0,1)$- Cauchy-matrix is the matrix consisting of 1's only.
- Cauchy matrices have applications in

  coding thery
  integral equations
  conformal mappings and
  singular integrals.
- **Problem:** Find an interpretation of $per(A)$ for Cauchy matrices A.

Computing $per(A) \cdot det(A)$, VII

Theorem A:(Carlitz and Levine, 1960)

Let $A = B^{(-1)}$ with B of rank 2 and for all i, j $b_{i,j} \neq 0$. Then again

$$per(A) \cdot det(A) = det(A^{(2)})$$

Lemma:
The rank of $B = A(\bar{x}, \bar{y})^{(-1)}$ is at most 2.

Hence Borchardt's theorem follows from Theorem A.

Computing $per(A) \cdot det(A)$, VIII

For a square matrix X and $\sigma \in \mathcal{S}_n$ we denote by X_σ the matrix with the rows permuted by σ.

Theorem B:(Sir Thomas Muir, 1882)

$$per(A) \cdot det(X) =$$

$$= \sum_{\sigma \in \mathcal{S}_n} (-1)^{sign(\sigma)} det(A \circ X_\sigma) \tag{3.6}$$

Each side contains $(n!)^2$ monomials, one for each pair of permutations (σ, τ).

Proof of Theorem B

We write for the left hand side

$$per(A) \cdot det(X) =$$

$$= \sum_{\phi, \psi \in \mathcal{S}_n} (-1)^{sign(\psi)} \prod_i a_{i,\phi(i)} \prod_i x_{i,\psi(i)}$$

We write for the right hand side

$$\sum_{\sigma \in \mathcal{S}_n} (-1)^{sign(\sigma)} det(A \circ X_\sigma)$$

For each monomial we have

$$(-1)^{sign(\sigma)} \left((-1)^{sign(\phi)} \prod_i (A \circ X_\sigma)_{i,\phi(i)} \right) =$$

$$(-1)^{sign(\sigma\phi)} \prod_i a_{i,\phi(i)} x_{\sigma(i),\phi(i)} =$$

$$(-1)^{sign(\phi\sigma^{-1})} \prod_i a_{i,\phi(i)} \prod_j x_{j,\phi\sigma^{-1}(j)}$$

Now we use that $\psi \leftrightarrow \phi\sigma^{-1}$ gives a bijection of permutations (for fixed σ).

Proof of Theorem A, I

Setting $X = A$ we get from Theorem B

$$per(A) \cdot det(A) =$$

$$= \sum_{\sigma \in \mathcal{S}_n} (-1)^{sign(\sigma)} det(A \circ A_\sigma)$$

It suffices to to show that for each $\sigma \neq Id_n$

$$det(A \circ A_\sigma) = 0$$

Then

$$per(A) \cdot det(A) = det(A \circ A) = det(A^{(2)})$$

Proof of Theorem A, II

If $\sigma \neq Id_n$ then $\sigma = \prod_k \tau_k$ is the product of disjoint cycles.

If one τ_{k_1} is a transposition, then $A \circ A_{\tau_{k_1}}$ has two identical rows, and hence $det(A \circ A_{\tau_{k_1}}) = 0$.

W.l.o.g., we can assume that each τ_k is a cycle of at least three elements.

We want to use that $B = A^{(-1)}$ is of rank at most 2.

Proof of Theorem A, III

Assume, w.l.o.g., that in $B = (b_{s,j})$ all the rows are linear combinations of the first two rows.

$$b_{s,j} = \lambda_s b_{1,j} + \mu_s b_{2,j}$$

Let $\sigma = \prod_k \tau_k$ with $\tau_1 = (1, 2, \ldots, h)$
and $3 \leq h \leq n$.

Let $C = (c_{i,j})$ be the $(h \times n)$ matrix consisting of the first h rows of $B^{(-1)} \circ B_\sigma^{(-1)} = A \circ A_\sigma$

$$c_{i,j} = b_{i,j}^{-1} b_{i+1,j}^{-1}$$

with addition $\mod h$.

Now put $K = Diag(k_1, \ldots, k_n)$ with $k_j = \prod_{s=1}^h b_{s,j}$. and $D = (d_{i,j}) = CK$. Hence

$$d_{i,j} = c_{i,j} \prod_s b_{s,j} = b_{i,j}^{-1} b_{i+1,j}^{-1} \prod_s b_{s,j} = \prod_{s \neq i, i+1} b_{s,j}$$

and $rank(C) = rank(D)$.

Proof of Theorem A, IV

Using

$$b_{s,j} = \lambda_s b_{1,j} + \mu_s b_{2,j}$$

we get

$$d_{i,j} = \prod_{s \neq i, i+1} \lambda_s b_{1,j} + \mu_s b_{2,j} =$$

$$\sum_{t=0}^{h-2} \nu_{i,t} b_{1,j}^{h-2-t} b_{2,j}^t$$

hence $rank(D) = rank(C) \leq h - 1$, and hence

$$det(A \circ A_\sigma) = 0$$

THIS COMPLETES THE PROOF

Computing $per(A(\bar{x}, \bar{y}))$

We look again at $A(\bar{x}, \bar{y})$ with $a_{i,j} = (x_i - y_j)^{-1}$

Assume

- $a_{i,j}^{-1} = (x_i - y_j) \neq 0$ for all $i, j \leq n$ and
- $det(A) \neq 0$

Then $per(A)$ can be computed in polynomial time by

$$per(A) = \frac{det(A^{(2)})}{det(A)}$$

What do we do when $det(A) = 0$?

The method of perturbation

Idea: Find a matrix A_ϵ such that

- A_ϵ is a Cauchy matrix
- $per(A_\epsilon) \cdot det(A_\epsilon)$ is a polynomial in ϵ of degree n
- $det(A_\epsilon) \not\equiv 0$

Compute

$$per(A_\epsilon) \cdot det(A_\epsilon) = det(A_\epsilon^{(2)})$$

for $n+1$ many values ϵ_i and interpolate.

May be tricky !

Barvinok's Theorems

1996

Theorem:
There as a function $g_{per}(r)$ such that for $(n \times n)$ matrices M over $\mathbb{Z}$ of rank r the **permanent** of M can be computed using $g_{per}(r) \cdot O(n^b)$ arithmetic operations for some constant $b = O(r)$.

Theorem:
There as a function $g_{ham}(r)$ such that for $(n \times n)$ matrices M over $\mathbb{Z}$ of rank r the **hamiltonian** of M can be computed using $g_{ham}(r) \cdot O(n^c)$ arithmetic operations for some constant $c = O(r^2)$.

TITLE

TEXT

TITLE

TEXT

TITLE

TEXT

TITLE

TEXT