

Problem 1: Determinants can be floppy

Problem Statement

Consider a non-singular matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, given by

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad (1)$$

Let us compute its determinant using the following formula:

$$\det(\mathbf{A}) := \sum_{k=1}^n (-1)^k a_{1k} \det(\hat{\mathbf{A}}^{1k})$$

where $\hat{\mathbf{A}}^{1k} \in \mathbb{R}^{(n-1) \times (n-1)}$ is the matrix obtained by removing the first row and k -th column of \mathbf{A} .

a

Suppose for $\mathbf{A} \in \mathbb{R}^{n \times n}$, computing $\det(\mathbf{A})$ using the above formula takes y_n flops. Find the relationship between y_n and y_{n-1} . Assume that negating a number does not require a flop.

b

Use your recurrence relation in (a) to prove that $n! \leq y_n \leq \frac{1}{2}(n+1)!$ for each $n \geq 5$. Here, $z! = z(z-1)(z-2)\dots 1$ denotes the factorial of z .

c

Estimate how much time your computer takes for one flop. Please submit your code.

d

- Using your answer in (c), estimate how long it will take to compute $\det(\mathbf{A})$ for the provided equation, using $n = 100$. Answer in number of years! The following Sterling's approximation may be useful:

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

- Based on your answer in (i), will you use this method to compute the determinant 100×100 matrix? By the way, the age of the universe is 13.8 billion years.

Solution

a

Let $F(n) := y_n$, i.e. the FLOP count of this calculation as a function of n . We can see that, for a given sub-term in the summation we have:

$$(-1)^k a_{1k} \det(\hat{\mathbf{A}}^{1k})$$

The sub-matrix determinant will take $F(n-1)$ FLOPs. The negation will take none. The multiplication of a_{1k} and the sub-determinant will take one. Each sub-term will thus take $F(n-1) + 1$ FLOPS, and there are n sub-terms. Then, the summations will require $n-1$ additional flops. So, we find:

$$F(n) = n(F(n-1) + 1) + (n-1) = nF(n-1) + 2n - 1$$

Note that $F(0) = F(1) = 0$, so:

n	1	2	3	4	5	6	7	8	9
$F(n)$	0	3	14	63	324	1955	13698	109599	986408

b

$$\begin{aligned}
 F(n) &= nF(n-1) + 2n - 1 \\
 &= n(n-1)F(n-2) + n(n-1) \\
 &= n((n-1)((n-1)(\dots) + n-2) + n-2) + n-1 \\
 &= 2n-1 + n(2(n-1)-1) + n(n-1)(2(n-2)-1) + \dots + n! \\
 &< 2n + 2n(n-1) + 2n(n-1)(n-2) + \dots + 2n! + n! \\
 &= 2(n + n(n-1) + \dots + n!/2) + 3n! \\
 &< 2 \left(\frac{n!}{(n-1)!} + \frac{n!}{(n-2)!} + \dots + \frac{n!}{2!} \right) + 3n!
 \end{aligned}$$

Then, note that:

$$\frac{2n!}{k!} \leq \frac{n!}{(k-1)!} \forall k \geq 2$$

and thus:

$$\begin{aligned}
 &\frac{n!}{(n-1)!} + \frac{n!}{(n-2)!} + \frac{n!}{(n-3)!} + \dots + \frac{n!}{2!} \\
 &\leq \frac{n!}{(n-2)!} + \frac{n!}{(n-2)!} + \frac{n!}{(n-3)!} + \dots + \frac{n!}{2!} \\
 &\leq \frac{n!}{(n-3)!} + \frac{n!}{(n-3)!} + \dots + \frac{n!}{2!} \\
 &\leq \frac{n!}{(n-4)!} + \dots + \frac{n!}{2!} \\
 &\quad \dots \\
 &\leq \frac{n!}{2!} + \frac{n!}{2!} \\
 &\leq n!
 \end{aligned}$$

So:

$$\begin{aligned}
 F(n) &= nF(n-1) + 2n - 1 < 2 \left(\frac{n!}{(n-1)!} + \frac{n!}{(n-2)!} + \dots + \frac{n!}{2!} \right) + 3n! \\
 &\leq 2n! + 3n! = 5n!
 \end{aligned}$$

Furthermore, note that for $n \geq 5$:

$$5n! \leq \frac{1}{2}(n+1)!$$

as

Referenced [1] for derivation.¹

c

¹There's something amusing about a recursive algorithm post on a website called "StackOverflow" ...

Problem 2: It Schur is Nice

Problem Statement

Now, let's compute the determinant of $\mathbf{A} \in \mathbb{R}^{n \times n}$ differently. Write \mathbf{A} as

$$\mathbf{A} = \begin{pmatrix} a_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \quad (2)$$

where $\mathbf{A}_{22} \in \mathbb{R}^{(n-1) \times (n-1)}$. Other blocks have appropriate dimensions. If $a_{11} \neq 0$, then

$$\det(\mathbf{A}) = a_{11} \det(\mathbf{A}_{22} - \mathbf{A}_{21}a_{11}^{-1}\mathbf{A}_{12})$$

where $\mathbf{A}_{22} - \mathbf{A}_{21}a_{11}^{-1}\mathbf{A}_{12}$ is non-singular. Suppose you can swap two rows without any FLOPs, but such a swap changes the sign of the determinant.

a

If computing $\det A$ using (2) takes z_n flops, find a recurrence relation between z_n and z_{n-1} .

b

Find the dominant term in n in the expression of z_n as a function of n ; i.e. without a recurrence relation.

c

Will you use (2) to compute the determinant of a 100×100 matrix? Answer logically.

Solution

a

Applying the scalar division to one of the matrices will require $n - 1$ operations. The matrix-matrix multiplication is really the outer product of a $(n - 1) \times 1$ and $1 \times (n - 1)$ vector, requiring $(n - 1) \times (n - 1) = (n - 1)^2$ operations. Finally, the subtraction of Matrices will require $(n - 1)^2$ operations. The result will be $n(n - 1)^2 + (n - 1)^2 = (n + 1)(n - 1)^2$ to produce an argument to the next determinant, and the result must be multiplied once more.

$$F(n) = F(n - 1) + (n + 1)(n - 1)^2 + 1$$

b

Ignoring the $F(n - 1)$ term, the resulting expression becomes:

$$F(n) = (n + 1)(n - 1)^2 + 1 = n^3 - n^2 - n + 2$$

so the algorithm is $\mathcal{O}(n^3)$.

c

Sure, for a 100×100 matrix, $n = 100 \implies n^3 = 1,000,000$. Given that the bulk of the expense are matrix-matrix operations (which are indeed highly parallelizable), I'd feel comfortable quoting the "theoretical" maximum FLOPS performance of my GPU, an nVidia RTX 2070 Super at 283.2 GFLOPS for FP64. This matrix would require $\approx \frac{100^3}{283.2 \cdot 10^9} = 3.53$ microseconds to calculate.

Of course, this algorithm isn't perfectly parallelizable, but even an order of magnitude slower would be 3.53 milliseconds...I'm probably not worried about that unless I'm finding the determinants of LOTS of these matrices.

Problem 3: A Plethora of Inversion Methods

Problem Statement

Suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ is nonsingular. Then, the inverse of \mathbf{A} can be computed elementwise as

$$[\mathbf{A}^{-1}]_{ij} = (-1)^{i+j} \frac{\det(\mathbf{B}_{ij})}{\det \mathbf{A}}$$

where \mathbf{B}_{ij} is the matrix obtained by removing j -th row and the i -th column of \mathbf{A} . Count the number of FLOPs (only dominant terms of n) required to compute the inverse of \mathbf{A} using the above equation. For computing determinants, assume that you are using the scheme in problem 2. Compare it with the number of FLOPs you will take to compute an LU decomposition followed by forward and backward substitutions to compute \mathbf{A}^{-1} .

Hint: To compute \mathbf{A}^{-1} using LU decompositions and forward/backward substitutions, you can calculate the columns x_1, \dots, x_n of \mathbf{A}^{-1} by solving $\mathbf{A}x_i = \mathbf{e}_i$ where $\mathbf{e}_i \in \mathbb{R}^n$ is a vector whose elements are all zeros except its i -th element that is unity.

Solution

Bibliography

- [1] AbcAeffchen. July 2014. URL: <https://stackoverflow.com/a/24793341>.