Dynamic Programming Practice

Matrix Multiplication

Find the best order of multiplication (or equivalently, the least costly parenthesization) of the product $A \cdot B \cdot C \cdot D \cdot E \cdot F$ with A an 5×8 matrix, B an 8×7 matrix, C an 7×12 matrix, D a 12×6 matrix, E a 6×20 matrix, and F a 20×10 matrix.

Solution

Two matrices

$A \cdot B$: $5 \cdot 8 \cdot 7 = 280$ multiplications
$B \cdot C$: $8 \cdot 7 \cdot 12 = 672$ multiplications
$C \cdot D$: $7 \cdot 12 \cdot 6 = 504$ multiplications
$D \cdot E$: $12 \cdot 6 \cdot 20 = 1440$ multiplications
$E \cdot F$: $6 \cdot 20 \cdot 10 = 1200$ multiplications

Three matrices

ABC: $(AB) \cdot C$ has $280 + 5 \cdot 7 \cdot 12 = 700$ multiplications <—
because AB is a 5×7 matrix and C a 7×12 matrix.
$A \cdot (BC)$ has $5 \cdot 8 \cdot 12 + 672 = 1152$ multiplications

BCD: $(BC) \cdot D$ has $8 \cdot 12 \cdot 6 + 672 = 1248$ multiplications
$B \cdot (CD)$ has $8 \cdot 7 \cdot 6 + 504 = 840$ multiplications <—

CDE: $(CD) \cdot E$ has $7 \cdot 6 \cdot 20 + 504 = 1344$ multiplications <—
$C \cdot (DE)$ has $7 \cdot 12 \cdot 20 + 1440 = 3120$ multiplications

DEF: $(DE) \cdot F$ has $12 \cdot 20 \cdot 10 + 1440 = 3840$ multiplications
$D \cdot (EF)$ has $12 \cdot 6 \cdot 10 + 1200 = 1920$ multiplications <—

We mark the best option with an arrow.

Four matrices

$ABCD$: $A \cdot (BCD)$ has $5 \cdot 8 \cdot 6 + 840 = 1080$ multiplications
$(AB) \cdot (CD)$ has $5 \cdot 7 \cdot 6 + 280 + 504 = 994$ multiplications, <—
since the first product costs 280, the second product costs 504 and AB is a 5×7 matrix and CD is a 7×6 matrix

$(ABC) \cdot D$ has $700 + 5 \cdot 12 \cdot 6 = 1060$ multiplications

$BCDE$: $B \cdot (CDE)$ has $8 \cdot 7 \cdot 20 + 1344 = 2464$ multiplications
$(BC) \cdot (DE)$ has $8 \cdot 12 \cdot 20 + 672 + 1440 = 4032$ multiplications
$(BCD) \cdot E$ has $8 \cdot 6 \cdot 20 + 840 = 1800$ multiplications <—

$CDEF$: $C \cdot (DEF)$ has $7 \cdot 12 \cdot 10 + 1920 = 2760$ multiplications
$(CD) \cdot (EF)$ has $7 \cdot 6 \cdot 10 + 504 + 1200 = 2124$ multiplications <—
$(CDE) \cdot F$ has $7 \cdot 20 \cdot 10 + 1344 = 2744$ multiplications because CDE is a 7×20 matrix, F is a 20×10 matrix and the best way to multiply CDE takes 1344 multiplications.

Five matrices

$ABCDE$: $A \cdot (BCDE)$ has $5 \cdot 8 \cdot 20 + 1800 = 2600$ multiplications
$(AB) \cdot (CDE)$ has $5 \cdot 7 \cdot 20 + 280 + 1344 = 2324$ multiplications
$(ABC) \cdot (DE)$ has $5 \cdot 12 \cdot 20 + 700 + 1440 = 3340$ multiplications
$(ABCD) \cdot E$ has $5 \cdot 6 \cdot 20 + 994 = 1594$ multiplications <—

$BCDEF$: $B \cdot (CDEF)$ has $8 \cdot 7 \cdot 10 + 2124 = 2684$ multiplications
$(BC) \cdot (DEF)$ has $8 \cdot 12 \cdot 10 + 672 + 1920 = 3552$ multiplications
$(BCD) \cdot (EF)$ has $8 \cdot 6 \cdot 10 + 840 + 1200 = 2520$ multiplications <—
$(BCDE) \cdot F$ has $8 \cdot 20 \cdot 10 + 1800 = 3400$ multiplications

---

Six matrices

$ABCDEF$ $A \cdot (BCDEF)$ has $5 \cdot 8 \cdot 10 + 2520 = 2920$ multiplications
$(AB) \cdot (CDEF)$ has $5 \cdot 7 \cdot 10 + 280 + 2124 = 2754$ multiplications
$(ABC) \cdot (DEF)$ has $5 \cdot 12 \cdot 10 + 700 + 1920 = 3220$ multiplications
$(ABCD) \cdot (EF)$ has $5 \cdot 6 \cdot 10 + 994 + 1200 = 2494$ multiplications <—
$(ABCDE) \cdot F$ has $5 \cdot 20 \cdot 10 + 1594 = 2594$ multiplications

Reading backward, we parenthesize $ABCDEF = (ABCD) \cdot (EF) = \big((AB)(CD)\big) \cdot (EF)$.

Levenshtein Distance between Strings

Find the Levenshtein distance and the minimum edits between ATAGGCT and GTGGAT.

Answer:

First step: We initialize the tableau.

		A	T	A	G	G	C	T
	0	1	2	3	4	5	6	7
G	1							
T	2							
G	3							
G	4							
A	5							
T	6							

We then fill in the tableau. At each step, the coefficient is the minimum of: The coefficient above plus one, the coefficient to the left plus 1, the coefficient up-and-to-the-left plus 1 if the letters are different and just the coefficient up-and-to-the-left if the letters are the same. This corresponds to dropping a letter, adding a letter, changing a letter and copying a letter. We also write down how we obtained the current coefficient. We use l for left, u for upper, and d for diagonal. If there is a tie, we write down both in order to make it easier to compare results.

		A	T	A	G	G	C	T
	0	1	2	3	4	5	6	7
G	1	1,d	2,dl	3,dl	3,d	4,dl	5,l	6,l
T	2	2,du	1,d	2,l	3,l	4,dl	5,dl	5,d
G	3	3,du	2,u	2,d	2,d	3,dl	4,l	5,l
G	4	4,du	3,u	3,du	3,du	2,d	3,l	4,l
A	5	4,d	4,u	3,d	4,dlu	3,u	3,d	4,dl
T	6	5,u	4,d	4,u	4,d	4,u	4,du	3,d

The lower right coefficient tells us the edit distance is 3. We reconstruct the changes backward. We get

- Copy T
- Change A to C
- Copy G
- Copy G
- Drop A
- Copy T
- Change A to G

Reversing the order, we get
ATAGGCT —> GTAGGCT —> GTGGCT —> GTGGAT —> GTGGCT
where of course copying steps are not visible.