1 Αλυσιδωτός πολλαπλασιασμός μητρών

Είναι γνωστό ότι το γινόμενο $A \times B$ δύο μητρών A και B με διαστάσεις $m \times n$ και $q \times p$ ορίζεται μόνο όταν $n = q$, δηλαδή όταν ο αριθμός n των στηλών της A ισούται με τον αριθμό q των γραμμών της B. Η μήτρα $A \times B$ που προκύπτει έχει διαστάσεις $m \times p$.

Για παράδειγμα, αν η A έχει διαστάσεις 3×7 και η B έχει διαστάσεις 7×5, τότε ορίζεται η μήτρα $A \times B$ με διαστάσεις 3×5, ενώ δεν ορίζεται το γινόμενο $B \times A$.

Επίσης, στην περίπτωση που ορίζονται τα γινόμενα $A \times B$ και $B \times A$, εν γένει, ισχύει ότι $A \times B \neq B \times A$.

Επιπλέον, είναι γνωστό ότι ο υπολογισμός του γινόμενου $A \times B$ δύο μητρών A και B με διαστάσεις $m \times n$ και $n \times p$ απαιτεί $m \cdot n \cdot p$ πολλαπλασιασμούς.

Για παράδειγμα, αν η A έχει διαστάσεις 3×7 και η B έχει διαστάσεις 7×5, τότε ο υπολογισμός του γινομένου $A \times B$ απαιτεί $3 \cdot 7 \cdot 5 = 105$ πολλαπλασιασμούς.

Επιπρόσθετα, η πράξη του πολλαπλασιασμού μητρών ικανοποιεί την προσεταιριστική ιδιότητα, δηλαδή αν A, B και C είναι μήτρες με διαστάσεις $m_0 \times m_1$, $m_1 \times m_2$ και $m_2 \times m_3$ τότε ισχύει ότι

$$(A \times B) \times C = A \times (B \times C),$$

και η μήτρα $A \times B \times C$ που προκύπτει έχει διαστάσεις $m_0 \times m_3$.

Για παράδειγμα, έστω οι μήτρες A, B και C με διαστάσεις 3×7, 7×5 και 5×4, τότε το γινόμενο $A \times B \times C$ είναι μία μήτρα με διαστάσεις 3×4

Παρατήρηση. Ο υπολογισμός του γινομένου $A \times B \times C$ μπορεί να γίνει με δύο τρόπους:

1ος $(A \times B) \times C$.

Δηλαδή, υπολογίζουμε πρώτα το γινόμενο $A \times B$ και στη συνέχεια το γινόμενο $(A \times B) \times C$.

Επειδή η A έχει διαστάσεις 3×7 και η B έχει διαστάσεις 7×5, ο υπολογισμός του γινομένου $A \times B$ απαιτεί $3 \times 7 \times 5 = 105$ πράξεις και προκύπτει μία μήτρα με διαστάσεις 3×5.

Επειδή η $A \times B$ έχει διαστάσεις 3×5 και η C έχει διαστάσεις 5×4, ο υπολογισμός του γινομένου $(A \times B) \times C$ απαιτεί $3 \times 5 \times 4 = 60$ πολλαπλασιασμούς.

Συνολικά, απαιτούνται $105 + 60 = 165$ πολλαπλασιασμοί.

2ος $A \times (B \times C)$.

Δηλαδή, υπολογίζουμε πρώτα το γινόμενο $B \times C$ και στη συνέχεια το γινόμενο $A \times (B \times C)$.

Επειδή η B έχει διαστάσεις 7×5 και η C έχει διαστάσεις 5×4, ο υπολογισμός του γινομένου $B \times C$ απαιτεί $7 \times 5 \times 4 = 140$ πράξεις και προκύπτει μία μήτρα με διαστάσεις 7×4.

Επειδή η A έχει διαστάσεις 3×7 και η $B \times C$ έχει διαστάσεις 7×4, ο υπολογισμός του γινομένου $A \times (B \times C)$ απαιτεί $3 \times 7 \times 4 = 84$ πολλαπλασιασμοί.

Συνολικά, απαιτούνται $140 + 84 = 224$ πολλαπλασιασμοί.

Παρατηρούμε ότι ενώ το γινόμενο $A \times B \times C$ είναι το ίδιο ανεξάρτητα από την σειρά υπολογισμού του, εν τούτοις οι πράξεις (πολλαπλασιασμού) που απαιτούνται για τον υπολογισμό του εξαρτώνται από την σειρά εκτέλεσης των πολλαπλασιασμών. Στο παράδειγμα αυτό ο πρώτος τρόπος δίνει σε σχέση με τον δεύτερο μια μείωση του αριθμού των πράξεων κατά $\frac{224-165}{224} \cdot 100 = 24\%$.

Επομένως, στις εφαρμογές όπου εκτελείται πολλές φορές ο υπολογισμός του γινομένου $A \times B \times C$ για μήτρες με τις συγκεκριμένες διαστάσεις αλλά στις οποίες αλλάζουν οι τιμές τους, ο πρώτος τρόπος υπολογισμού δίνει μια οικονομία περίπου 24% στον απαιτούμενο χρόνο.

Στη συνέχεια θα μελετήσουμε το επόμενο γενικό πρόβλημα.

Πρόβλημα. Έστω οι μήτρες A_1, A_2, …, A_n με διαστάσεις $m_0 \times m_1$, $m_1 \times m_2$, …, $m_{n-1} \times m_n$. Να βρεθεί η σειρά υπολογισμού του γινομένου

$$A_1 \times A_2 \times \cdots \times A_n$$

η οποία ελαχιστοποιεί τις πράξεις (πολλαπλασιασμού) που απαιτούνται.

Μια πρώτη προσέγγιση σ' αυτό το πρόβλημα είναι η εύρεση των πράξεων που απαιτούνται για τον υπολογισμό του γινομένου με όλους τους δυνατούς τρόπους, και στη συνέχεια η επιλογή του τρόπου ο οποίος απαιτεί τον ελάχιστο αριθμό πράξεων.

Γι' αυτό θα μελετήσουμε πρώτα το πρόβλημα με πόσους διαφορετικούς τρόπους μπορεί να υπολογισθεί το γινόμενο

$$A_1 \times A_2 \times \cdots \times A_n$$

Ξεκινάμε με ειδικές περιπτώσεις:
Ο υπολογισμός του γινομένου $A_1 \times A_2$ γίνεται με 1 τρόπο.
Ο υπολογισμός του γινομένου $A_1 \times A_2 \times A_3$ γίνεται με 2 τρόπους:

$$(A_1 \times A_2) \times A_3$$

$$A_1 \times (A_2 \times A_3)$$

Ο υπολογισμός του γινομένου $A_1 \times A_2 \times A_3 \times A_4$ γίνεται με 5 τρόπους:

$$((A_1 \times A_2) \times A_3) \times A_4$$
$$(A_1 \times A_2) \times (A_3 \times A_4)$$
$$(A_1 \times (A_2 \times A_3)) \times A_4$$
$$A_1 \times ((A_2 \times A_3) \times A_4)$$
$$A_1 \times (A_2 \times (A_3 \times A_4))$$

Έστω a_n ο αριθμός των τρόπων υπολογισμού του γινομένου $A_1 \times A_2 \cdots \times A_n$, όπου $n \geq 2$. Τότε, από τα προηγούμενα $a_2 = 1$ και $a_3 = 2$ και $a_4 = 5$. Επίσης, μπορούμε να συμφωνήσουμε ότι $a_1 = 1$, διότι όταν έχουμε μόνο μία μήτρα A_1 υπάρχει μόνο ένας τρόπος να τον θεωρήσουμε.

Για τον υπολογισμό του a_n παρατηρούμε σε κάθε τρόπο θα υπάρχει ο τελευταίος πολλαπλασιασμός για τον υπολογισμό του γινομένου

$$(A_1 \times A_2 \times \cdots \times A_{k-1}) \times (A_k \times \cdots \times A_n)$$

για κάποιο $k = 2, \ldots, n$.

Οι υπολογισμοί των γινομένων $(A_1 \times A_2 \times \cdots \times A_{k-1})$ και $(A_k \times \cdots \times A_n)$ γίνονται ανεξάρτητα. Στο πρώτο γινόμενο υπάρχουν $k-1$ μήτρες, επομένως ο αριθμός των τρόπων υπολογισμού του είναι a_{k-1}, ενώ στο δεύτερο γινόμενο υπάρχουν $n-k+1$ μήτρες, επομένως ο αριθμός των τρόπων υπολογισμού του είναι a_{n-k+1}. Από την πολλαπλασιαστική αρχή, έπεται ότι υπάρχουν $a_{k-1} \cdot a_{n-k+1}$ τρόποι υπολογισμού του γινομένου

$$(A_1 \times A_2 \times \cdots \times A_{k-1}) \times (A_k \times \cdots \times A_n)$$

Αθροίζοντας για όλες τις δυνατές τις τιμές του k, δηλαδή για $k = 2, \ldots, n$ προκύπτει ότι ο αριθμός των τρόπων υπολογισμού του γινομένου

$$A_1 \times A_2 \times \cdots \times A_n$$

ικανοποιεί την αναδρομική σχέση

$$a_n = \sum_{k=2}^{n} a_{k-1} a_{n-k+1}$$

Επομένως,

$$a_5 = \sum_{k=2}^{5} a_{k-1} \cdot a_{6-k} = a_1 \cdot a_4 + a_2 \cdot a_3 + a_3 \cdot a_2 + a_4 \cdot a_1 = 1 \cdot 5 + 1 \cdot 2 + 2 \cdot 1 + 5 \cdot 1 = 14$$

$$a_6 = \sum_{k=2}^{6} a_{k-1} \cdot a_{7-k} = a_1 \cdot a_5 + a_2 \cdot a_4 + a_3 \cdot a_3 + a_4 \cdot a_2 + a_5 \cdot a_1 = 1 \cdot 14 + 1 \cdot 5 + 2 \cdot 2 + 5 \cdot 1 + 14 \cdot 1 = 42$$

Μερικές από τις αρχικές τιμές τιμές της ακολουθίας a_n δίδονται στον επόμενο πίνακα:

n	1	2	3	4	5	6	7	8	9	10	11	12	13
a_n	1	1	2	5	14	42	132	429	1430	4862	16796	58786	208012

Εύκολα προκύπτει ότι η ακολουθία a_n ισούται με τον $(n-1)$-οστό αριθμό Catalan C_{n-1}, και ισχύουν ότι

$$a_{n+1} = \frac{1}{n+1}\binom{2n}{n}$$

και

$$a_n = O(4^n)$$

Επομένως, ο αριθμός των διαφορετικών τρόπων υπολογισμού του γινομένου

$$A_1 \times A_2 \times \cdots \times A_n$$

αυξάνει εκθετικά και άρα δεν είναι εφικτός ο εξαντλητικός έλεγχος των πράξεων που απαιτούνται σε όλους τους δυνατούς τρόπους υπολογισμού του γινομένου αυτού, εκτός από τις περιπτώσεις με 2 ή 3 πίνακες.

Στη συνέχεια θα παρουσιάσουμε έναν αλγόριθμο υπολογισμού του ελάχιστου αριθμού των πράξεων ο οποίος βασίζεται στην εύρεση των κατάλληλων υποπροβλημάτων και στον ‘‘προγραμματισμό’’ της σειράς με την οποία πρέπει να λυθούν.

Όπως είδαμε, για τον υπολογισμό του γινομένου

$$A_1 \times A_2 \times \cdots \times A_n$$

για τις μήτρες A_1, A_2, ..., A_n με διαστάσεις $m_0 \times m_1$, $m_1 \times m_2$, ..., $m_{n-1} \times m_n$, υπάρχει ο τελευταίος πολλαπλασιασμός

$$(A_1 \times A_2 \times \cdots \times A_{k-1}) \times (A_k \times \cdots \times A_n)$$

για κάποιο $k = 2, \ldots, n$.

Η βασική ιδέα είναι ότι στη βέλτιστη λύση έχουμε σχηματίσει με βέλτιστο τρόπο τις τελευταίες μήτρες $A_1 \times A_2 \cdots A_{k-1}$ και $A_k \times A_{k+1} \cdots A_{n+1}$ και απομένει να εκτελέσουμε τον πολλαπλασιασμό τους.

Ο συνολικός ελάχιστος αριθμός πράξεων είναι το άθροισμα των πράξεων για τον υπολογισμό του γινομένου $A_1 \times A_2 \cdots A_{k-1}$ και του γινομένου $A_k \times A_{k+1} \cdots A_n$ συν το γινόμενο $m_0 \cdot m_{k-1} \cdot m_n$.

Επομένως, προκειμένου να βρούμε τον ελάχιστο αριθμό πράξεων θα εξετάσουμε την βέλτιστη λύση των υποπροβλημάτων $A_i \times A_{i+1} \times \cdots \times A_j$.

Η σειρά με την υποία θα λύσουμε τα υποπροβλήματα $A_i \times A_{i+1} \times \cdots \times A_j$ είναι ξεκινώντας από υποπροβλήματα με 1 όρο, στη συνέχεια με 2 όρους, μετά με 3 όρους, κ.ο.κ. μέχρι να φτάσουμε στο ζητούμενο γινόμενο με n όρους.

Έστω $C(i,j)$ ο ελάχιστος αριθμός των πράξεων για τον υπολογισμό του γινομένου $A_i \times A_{i+1} \times \cdots \times A_j$

Εύκολα προκύπτει ότι

Αν $j = i$, τότε

$$C(i,i) = 0.$$

Αν $j > i$, τότε

$$C(i,j) = \min\{C(i,k) + C(k+1,j) + m_{i-1} \cdot m_k \cdot m_j\}$$

για όλες τις τιμές του k με $i \leq k < j$.

Η απάντηση που ζητάμε είναι η τιμή του $C(1,n)$.

Παράδειγμα. Να βρεθεί με ποιο τρόπο να υπολογισθεί το γινομενο $A_1 \times A_2 \times A_3 \times A_4$, για τις μήτρες A_1, A_2, A_3 και A_4 με διαστάσεις 5×2, 2×6, 6×4 και 4×10, έτσι ώστε ο αριθμός των πράξεων να είναι ελάχιστος.

Απάντηση. Θα λύσουμε όλα τα υποπροβλήματα $A_i \times A_{i+1} \times \cdots \times A_j$ όπου $1 \leq i \leq j \leq 4$.

Έστω $C(i,j)$ ο ελάχιστος αριθμός των πράξεων για τον υπολογισμό του γινομένου $A_i \times A_{i+1} \times \cdots \times A_j$.

Θα ξεκινήσουμε τον υπολογισμό από υποπροβλήματα με 1 όρο. Έχουμε τα υποπροβλήματα A_1, A_2, A_3 και A_4, για τα οποία ισχύει ότι

$$C(1,1) = C(2,2) = C(3,3) = C(4,4) = 0$$

Με δύο όρους, έχουμε τα υποπροβλήματα $A_1 \times A_2$, $A_2 \times A_3$ και $A_3 \times A_4$, για τα οποία ισχύει ότι

$$C(1,2) = 5 \cdot 2 \cdot 6 = 60$$

$$C(2,3) = 2 \cdot 6 \cdot 4 = 48$$

$$C(3,4) = 6 \cdot 4 \cdot 10 = 240$$

Με τρεις όρους, έχουμε τα υποπροβλήματα $A_1 \times A_2 \times A_3$ και $A_2 \times A_3 \times A_4$, για τα οποία ισχύει ότι

$$\begin{aligned}
C(1,3) &= \min\{C(1,k) + C(k+1,3) + m_0 \cdot m_k \cdot m_3\} \text{ όπου } 1 \leq k < 3 \\
&= \min\{C(1,1) + C(2,3) + m_0 \cdot m_1 \cdot m_3, C(1,2) + C(3,3) + m_0 \cdot m_2 \cdot m_3\} \\
&= \min\{0 + 48 + 5 \cdot 2 \cdot 4, 60 + 0 + 5 \cdot 6 \cdot 4\} \\
&= \min\{48 + 40, 60 + 120\} \\
&= 88.
\end{aligned}$$

η βέλτιστη λύση στην περίπτωση αυτή προκύπτει από τον τρόπο υπολογισμού $A_1 \times (A_2 \times A_3)$, (διότι $C(1,3) = C(1,1) + C(2,3) + m_0 \cdot m_1 \cdot m_3$ δηλαδή υπολογίζουμε το γινόμενο $A_2 \times A_3$ και στη συνέχεια το αποτέλεσμα επί A_1).

$$\begin{aligned}
C(2,4) &= \min\{C(2,k) + C(k+1,4) + m_1 \cdot m_k \cdot m_4\} \text{ όπου } 2 \leq k < 4 \\
&= \min\{C(2,2) + C(3,4) + m_1 \cdot m_2 \cdot m_4, C(2,3) + C(4,4) + m_1 \cdot m_3 \cdot m_4\} \\
&= \min\{0 + 240 + 2 \cdot 6 \cdot 10, 48 + 0 + 2 \cdot 4 \cdot 10\} \\
&= \min\{240 + 120, 48 + 80\} \\
&= 128.
\end{aligned}$$

η βέλτιστη λύση στην περίπτωση αυτή προκύπτει από τον τρόπο υπολογισμού $(A_2 \times A_3) \times A_4$, (διότι $C(2,4) = C(2,3) + C(4,4) + m_1 \cdot m_3 \cdot m_4$ δηλαδή υπολογίζουμε το γινόμενο $A_2 \times A_3$ και στη συνέχεια το αποτέλεσμα επί A_4).

Τέλος, με τέσσερεις όρους, έχουμε το πρόβλημα $A_1 \cdot A_2 \cdot A_3 \cdot A_4$, για το οποίο ισχύει ότι

$$\begin{aligned}
C(1,4) &= \min\{C(1,k) + C(k+1,4) + m_0 \cdot m_k \cdot m_4\} \text{ όπου } 1 \leq k < 4 \\
&= \min\{C(1,1) + C(2,4) + m_0 \cdot m_1 \cdot m_4, C(1,2) + C(3,4) + m_0 \cdot m_2 \cdot m_4, C(1,3) + C(4,4) + m_0 \cdot m_3 \cdot m_4\} \\
&= \min\{0 + 128 + 5 \cdot 2 \cdot 10, 60 + 240 + 2 \cdot 6 \cdot 10, 88 + 0 + 5 \cdot 4 \cdot 10\} \\
&= \min\{128 + 100, 60 + 240 + 120, 88 + 200\} \\
&= 228
\end{aligned}$$

η βέλτιστη λύση στην περίπτωση αυτή προκύπτει από τον συνδυασμό της μήτρας A_1 με τον βέλτιστο τρόπο υπολογισμού του γινομένου $A_2 \times A_3 \times A_4$, ο οποίο όπως υπολογίσαμε στο προηγούμενο βήμα ισούται με $(A_2 \times A_3) \times A_4$.

Επομένως, η σειρά υπολογισμού του γινομένου $A_1 \times A_2 \times A_3 \times A_4$ η οποία ελαχιστοποιεί τον αριθμό των πράξεων που απαιτούνται είναι

$$A_1 \times ((A_2 \times A_3) \times A_4)$$

με 228 πράξεις.

Η χειρότερη περίπτωση είναι ο υπολογισμός του γινομένου με τη σειρά $A_1 \times (A_2 \times (A_3 \times A_4))$, η οποία απαιτεί 460 πράξεις, δηλαδή το διπλάσιο αριθμό πράξεων από τη βέλτιστη περίπτωση.

Άσκηση. Να βρεθεί η σειρά υπολογισμού του γινομένου

$$A_1 \times A_2 \times A_3 \times A_4 \times A_5$$

η οποία ελαχιστοποιεί τις απαιτούμενες πράξεις, για τις μήτρες A_1, A_2, A_3, A_4 και A_5 με διαστάσεις

(α) 7×3, 3×5, 5×7, 7×4 και 4×5.

(β) 2×3, 3×4, 4×5, 5×6 και 6×7.

(γ) 2×6, 6×5, 5×5, 5×6 και 6×2.

(δ) 2×3, 3×2, 2×3, 3×2 και 2×3.

(ε) 2×4, 4×5, 5×3, 3×6 και 6×2.

Ποιος είναι ο χειρότερος τρόπος σε κάθε περίπτωση;