Mathematical Foundations for Machine Learning

BITS Pilani
Pilani Campus

MFML Team

innovate achieve lead

BITS Pilani
Pilani Campus

S2 -22 AIMLCZC416, MFML

Webinar#1

Agenda

Discussion on Solutions of

- Homework Problems
- Proof

Homework Problem

Q1 Let $A_{m\times n}$ be a given matrix with $m > n$. If the time taken to compute the determinant of a square matrix of size j is j^3, find upper bound on the

a) total time taken to find the rank of A using determinants

b) number of additions and multiplications required to determine the rank using the elimination procedure.

Homework Problem

a) total time taken to find the rank of A using determinants

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m \times n}$

Since m> n the maximum possible rank of A is n.

Thus, we need to compute all determinates of order k, 1<k<n.

Therefore, Number of Determinants = $\sum_{i=1}^{i=\min(m,n)} {}^{m}C_i * {}^{n}C_i$,

Where, $\min(m,n)$ is maximum size of square matrix in A

${}^{m}C_i$ is the combination of 'i' columns out of m

${}^{n}C_i$ is the combination of 'i' rows out of n

Given that time taken to compute the determinant of a square matrix of size i is i^3

Hence total time taken to compute determinant of a square matrix is $= \sum_{i=1}^{i=\min(m,n)} {}^{m}C_i \, {}^{n}C_i \, i^3$

Homework Problem

b) number of additions and multiplications required to determine the rank using the elimination procedure.

Solution:

Let $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m\times n}$

The number of multiplications required to make all the entries below a_{11} to zeros is $n(m-1)$

The number of multiplications required to make all the entries below a_{22} to zeros is $(n-1)(m-2)$

.

The number of multiplications required to make all the entries below a_{kk} to zeros is $(n-k+1)(m-k)$

Therefore, total number of multiplications required is $= n(m-1)+(n-1)(m-2)+\ldots$

$$= \sum_{K=1}^{n}(m-k)(n-k+1)$$

Similarly, total number of additions required is $= \sum_{K=1}^{n}(m-k)(n-k+1)$

Hence the total number additions and multiplications required $= 2\sum_{K=1}^{n}(m-k)(n-k+1)$

Homework Problem

Q2

Let A n×n be a given square matrix. Compute the number of multiplications and additions required to evaluate A^{28} using

a) the naive method, $A^{28} = (A \cdot A \cdots A)$28 times
b) A^2, $A^4 = A^2 \cdot A^2$, etc.

Solution:

Let
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}_{n\times n},$$

Now consider $A^2 = A * A$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}_{n\times n} * \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}_{n\times n}$$

Homework Problem

For multiplying the above matrices, we have

	Number of operations required for multiplication of R_i with all columns		Number of operations required for addition of R_i with all columns	
	Multiplications required	Number of operations required for multiplication of	Additions required	Number of operations required for addition of
$R_1C_1 = a_{11}a_{11} + a_{12}a_{21} + \ldots + a_{1n}a_{n1}$ $R_1C_2 = a_{11}a_{12} + a_{21}a_{22} + \ldots + a_{1n}a_{n2}$ $\vdots$ $R_1C_n = a_{11}a_{1n} + a_{12}a_{2n} + \ldots + a_{1n}a_{nn}$	n n $\vdots$ n	1st Row with all columns is n^2	$(n-1)$ $(n-1)$ $\vdots$ $(n-1)$	1st Row with all columns is $n(n-1)$
$R_2C_1 = a_{21}a_{11} + a_{22}a_{21} + \ldots + a_{2n}a_{n1}$ $R_2C_2 = a_{21}a_{12} + a_{22}a_{22} + \ldots + a_{2n}a_{n2}$ $\vdots$ $R_2C_n = a_{21}a_{1n} + a_{22}a_{2n} + \ldots + a_{2n}a_{nn}$	n n $\vdots$ n	2nd Row with all columns is n^2	$(n-1)$ $(n-1)$ $\vdots$ $(n-1)$	2nd Row with all columns is $n(n-1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$R_nC_1 = a_{n1}a_{11} + a_{n2}a_{21} + \ldots + a_{nn}a_{n1}$ $R_nC_2 = a_{n1}a_{12} + a_{n2}a_{22} + \ldots + a_{nn}a_{n2}$ $\vdots$ $R_nC_n = a_{n1}a_{1n} + a_{n2}a_{2n} + \ldots + a_{nn}a_{nn}$	n n $\vdots$ n	nth Row with all columns is n^2	$(n-1)$ $(n-1)$ $\vdots$ $(n-1)$	nth Row with all columns is $n(n-1)$
Total	$n\,n^2 = n^3$		$n\,n\,(n-1) = n^2(n-1)$	

Thus, the number of multiplications required to multiply the matrix A^2 is $= n^3$

Thus, the number of additions required to multiply the matrix A^2 is $= n^2(n-1)$

Homework Problem

a) For multiplication of the form the naive method, A^{28}= (A·A···A) 28 times is as fallows

The required number of additions and multiplications to get A^2 is $= n^3 + n^2(n-1)$
The required number of additions and multiplications to get A^3 is $= 2\left[n^3 + n^2(n-1)\right]$
And so on…

Similarly, the required number of additions and multiplications to get A^{28} is $= 27\left[n^3 + n^2(n-1)\right]$

b) For multiplication of the form A^2, A^4=A^2·A^2, etc.

We know that
The number of additions and multiplications to get A^2 is $= n^3 + n^2(n-1)$
The required number of additions and multiplications to get $A^4 = A^2 * A^2$ is $= n^3 + n^2(n-1)$
The required number of additions and multiplications to get $A^8 = A^4 * A^4$ is $= n^3 + n^2(n-1)$
The required number of additions and multiplications to get $A^{16} = A^8 * A^8$ is $= n^3 + n^2(n-1)$
The required number of additions and multiplications to get $A^{28} = A^{16} * A^8 * A^4$ is $= 2\left[n^3 + n^2(n-1)\right]$

Thus, the total no of operations required to get A^{28} is $= 6\left[n^3 + n^2(n-1)\right]$.

Homework Problem

Q3: Let A_{nxk} and B_{kxk} be two given matrices with rank(B) =k. Estimate the rank of AB in terms of the ranks of A and B. Note that the rank of a matrix is the dimension of the vector space spanned by its rows/ columns.

Consider the spaces

C_1 ={Ax : x is a vector of length k}

C_2 ={ABy : y is a vector of length k}

C_1 and C_2 are vector spaces. Let y be any vector of length k and x = By. Since ABy = Ax belongs to C_1

C_2 is a subspace of C_1

dim(C2) ≤ dim(C1)

rank (AB) ≤ rank(A)

Homework Problem

Applying to the transpose

$\text{rank}(AB)^t = \text{rank}(B^tA^t) \leq \text{rank}(B^t) = \text{rank}(B)$

$\text{Rank}\ (AB) \leq \text{rank}(B)$

Thus

$\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$

Homework Problem

Q4: Compute the total number of divisions, multiplications and additions required to perform the forward elimination and back substitution in solving a system of linear equations $A_{nxn}\, x = b$ using the Gauss elimination method.

Homework Problem

Number of divisions required in Gauss elimination $= \sum_{k=1}^{n-1}(n-k)$

Number of multiplications and subtractions in Gauss elimination $= 2\sum_{k=1}^{n-1}(n-k)(n-k+1)$

$\Bigg\}$ Forward elimination

Number of operations required back substitution $= 2\sum_{k=1}^{n}(n-k) + n$ $\Big\}$ Back substitution.

Now Let $f(n)$ be forward elimination $= \sum_{k=1}^{k=n-1}(n-k) + 2\sum_{k=1}^{k=(n-1)}(n-k)(n-k+1)$

$\sum_{k=1}^{n} k = 1+2+3+\cdots n = \frac{n(n+1)}{2}$

$\sum_{k=1}^{n} k^2 = \frac{k(k+1)(2k+1)}{6}$

Now $n-k=s$, $n-s=k$ | $k=1$, $n-s=1$, $s=n-1$ | $k=n-1$, $n-s=n-1$, $s=1$

Homework Problem

$$= \sum_{s=1}^{s=n-1} s + 2\sum_{s=1}^{s=n-1} s(s+1)$$

$$= \frac{(n-1)(n)}{2} + 2\left[\frac{(n-1)n(2n-1)}{6} + \frac{(n-1)n}{2}\right]$$

$$= \frac{n^2-n}{2} + \cancel{2}\left[\frac{(n^2-n)(2n-1) + 3n^2-3n}{\cancel{6}\,3}\right]$$

$$= \frac{3n^2-3n + 2(2n^3-2n^2-n^2+n+3n^2-3n)}{6}$$

$$= \frac{3n^2-3n + 2(2n^3-\cancel{3n^2}+n+\cancel{3n^2}-3n)}{6}$$

$$= \frac{3n^2-3n+4n^3+2n-6n}{6}$$

$$= \frac{4n^3}{6} + \frac{3n^2}{6} - \frac{7n}{6}$$

$$\boxed{f(n) = \frac{2}{3}n^3 + \frac{n^2}{2} - \frac{7}{6}n} \Rightarrow \{O(n^3) = f(n)$$

Homework Problem

Now

$$b(n) = 2\sum_{k=1}^{n}(n-k) + n$$

$$= 2\left(n * \sum_{k=1}^{n} 1 - \sum_{k=1}^{n} k\right) + n$$

$$= 2\left(n \cdot n - \frac{n(n+1)}{2}\right) + n$$

$$= 2\left(n^2 - \frac{n^2}{2} - \frac{n}{2}\right) + n$$

$$= 2\left(\frac{n^2}{2} - \frac{n}{2}\right) + n$$

$$= n^2 - \cancel{n} + \cancel{n}$$

$$\boxed{b(n) = n^2}$$

$$\{\, O(n^2) = b(n) \,\}$$

Homework Problem

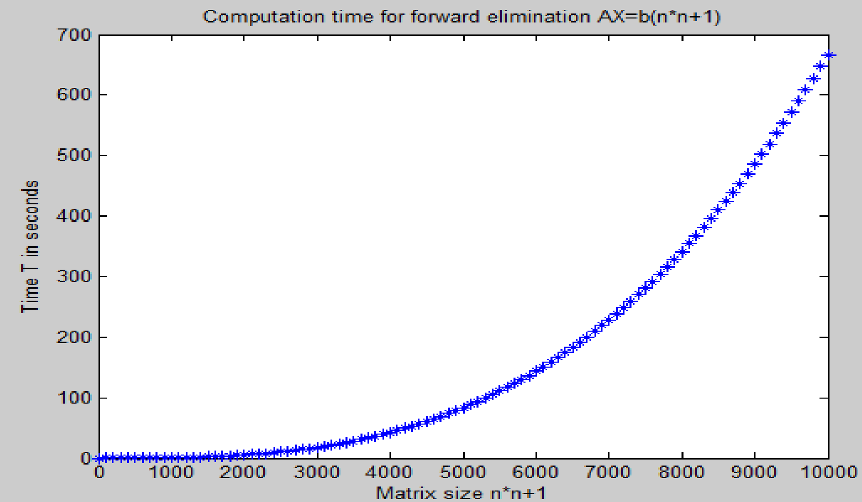
For solving $Ax = B$; $[A:B]_{(n \times n+1)}$

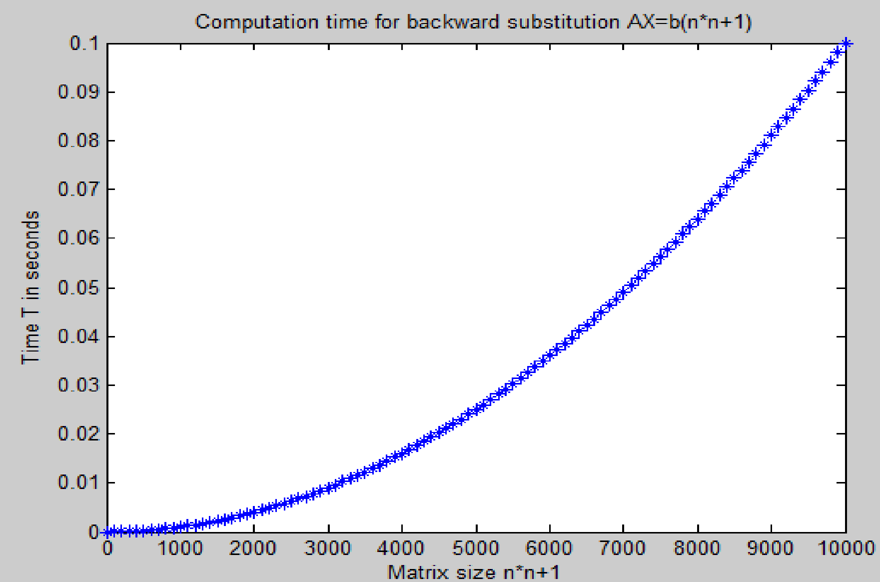
The number of operations required is

If for 1 count – 10^{-9} sec

Algorithm	$n = 1000$	$n = 10,000$
Elimination $f(n) = \frac{2}{3}n^3 + \frac{n^2}{2} - \frac{7}{6}n$	0.6671655 ≅ 0.7 Sec	666.7166 Sec ≅ 11.1119 min
Back Substitution $b(n) = n^2$	0.001 Sec	0.1 Sec

```
N=0:100:10000;
f=(2/3)*(N.^3)+(1/2)*(N.^2)-(7/6)*(N);
%f=(N.^2); for backward sub
T=f*10^(-9);
plot(N,T,'*')
```





Theorem: Let V be a finite dimensional vector space. Then any two bases have the same cardinality.

Proof: Suppose that $B=\{v_1,v_2,\ldots,v_m\}$ **and** $C=\{w_1,w_2,\ldots w_n\}$

Are two bases of V.

We want to show that m=n.

We may assume that m, n > 0.

Consider the set $B\cup\{w_n\}$ **.** $w_n\in V$ **Belongs to the span of**

$v_1, v_2, \ldots, v_m.$

Therefore, we may find $r_1, r_2, \ldots, r_m$ **such that**

$$w_n = r_1v_1 + r_2v_2 + \ldots + r_mv_m.$$

$w_n \neq 0$ and so at least one of $r_1, r_2, \ldots, r_m$ is non-zero.
We may assume that $r_m \neq 0$.
In this case

$$v_m = w_n - \left(r_1 v_1 + r_2 v_2 + \ldots + r_{m-1} v_{m-1}\right) / r_m$$

and so v_m is a linear combination of $v_1, v_2, \ldots, v_m$ and w_n.
We know that If v_n is a lnear combination of $v_1, v_2, \ldots, v_{n-1}$ then span$\{v_1, v_2, \ldots, v_{n-1}\}$=span$\{v_1, v_2, \ldots, v_n\}$
Therefore, $B - \{v_m\} \cup \{w_n\}$ also spans V.
So, now consider $v_1, v_2, \ldots, v_{m-1}, w_{n-1}$ and w_n.
As $v_1, v_2, \ldots, v_{m-1}$ and w_n span V it follows that w_{n-1} is a

linear combination of $v_1, v_2, \ldots, v_{m-1}$ and w_n.

$$w_{n-1} = r_1 v_1 + r_2 v_2 + \ldots + r_{m-1} v_{m-1} + s_n w_n$$

Suppose that every $r_i = 0$. Then w_{n-1} and w_n are dependent, which contradicts the fact that C is a basis. Thus $r_i \neq 0$ some i. Relabelling we may suppose that $r_{m-1} \neq 0$. As before this implies that v_{m-1} is a linear combination of $v_1, v_2, \ldots, v_{m-2}, w_{n-1}$ and w_n. But from the known results $v_1, v_2, \ldots, v_{m-1}, w_{n-1}$ and w_n span V.

We can repeat this process for every vector in C. It follows that m ≤ m. By symmetry n ≤ m. But then m=n.

Product of any two lower triangular matrices is a lower triangular matrix

A=[a_n], B=[b_n] be lower triangular matrices of order n and let C=AB

(i) The diagonal elements of C are given by

$$\forall j \in [1, 2, \ldots, n] : c_{jj} = a_{jj} b_{jj}$$

(ii) The matrix C is itself lower triangular matrix

Proof: From the definition of matrix product, we have

$$\forall i, j \in [1, 2, \ldots n],\ c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.\ When\, i = j,\ c_{jj} == \sum_{k=1}^{n} a_{jk} b_{kj}$$

Now both A and B are lower triangular matrices

Product of any two lower triangular matrices is a lower triangular matrix

Thus:

$$\left.\begin{array}{l} if\ k < j,\ b_{kj} = 0\ and\ a_{jk}b_{kj} = 0 \\ if\ k > j, a_{jk} = 0\ and\ a_{jk}b_{kj} = 0 \\ and\ a_{jk}b_{kj} \neq 0\ when\ j = k \end{array}\right\} \Rightarrow c_{jj} = \sum_{k=1}^{n} a_{jk}b_{kj} = a_{jj}b_{jj}$$

Now if $i < j$, it follows that either a_{ik} or b_{kj} is zero for all k, and thus $c_{ij} = 0$.

Therefore, C is also a lower triangular matrix.

Thank You