MATH3322 Matrix Computation
Homework 1

Due date: 22 February, Monday

1. Let
$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$
Find
$$\boldsymbol{A}\boldsymbol{x}, \quad \boldsymbol{A}^T\boldsymbol{y}, \quad \boldsymbol{A}\boldsymbol{B}^T.$$

2. Using only the definition of matrix transpose (i.e., the (i,j)-th entry of $\boldsymbol{A}^T$ is a_{ji}), prove rigorously the following transpose formula for block matrices
$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1q} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \dots & \boldsymbol{A}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{p1} & \boldsymbol{A}_{p2} & \dots & \boldsymbol{A}_{pq} \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{A}_{11}^T & \boldsymbol{A}_{21}^T & \dots & \boldsymbol{A}_{p1}^T \\ \boldsymbol{A}_{12}^T & \boldsymbol{A}_{22}^T & \dots & \boldsymbol{A}_{p2}^T \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{1q}^T & \boldsymbol{A}_{2q}^T & \dots & \boldsymbol{A}_{pq}^T \end{bmatrix},$$
where the block matrices on both sides of the equality are valid matrix partitions.

3. Compare performances of different implementations of the matrix-vector multiplication. An incomplete Matlab code `matvecprod.m` is attached.
   (a) Complete the code for matrix-vector multiplication.
   (b) Run the program on your computer and record the time needed for each implementation.

   Here are two remarks. 1. Different implementations may have very different performances, though mathematically they are equivalent. 2. We should call build-in functions for matrix computations if they are available, because they are optimized for your computer.

4. We have seen that a standard matrix multiplication $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$, where $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$, needs $2n^3$ flops (ignoring the lower order terms). This computational complexity can be reduced as in the following.
   (a) Suppose n is even. We decompose the matrix multiplication $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ into 2×2 block matrix multiplication
$$\begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix},$$
   where all sub-matrices are of size $\frac{n}{2} \times \frac{n}{2}$. In a standard matrix multiplication, we need to compute the blocks via $\boldsymbol{C}_{ij} = \boldsymbol{A}_{i1}\boldsymbol{B}_{1j} + \boldsymbol{A}_{i2}\boldsymbol{B}_{2j}$, for $i,j = 1,2$. How many sub-matrix multiplications and sub-matrix additions are needed for $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$?

(b) We may compute the block matrices via the following sub-matrix operations:

$$
\begin{aligned}
\boldsymbol{P}_1 &= (\boldsymbol{A}_{11} + \boldsymbol{A}_{22})(\boldsymbol{B}_{11} + \boldsymbol{B}_{22}),\\
\boldsymbol{P}_2 &= (\boldsymbol{A}_{21} + \boldsymbol{A}_{22})\boldsymbol{B}_{11},\\
\boldsymbol{P}_3 &= \boldsymbol{A}_{11}(\boldsymbol{B}_{12} - \boldsymbol{B}_{22}),\\
\boldsymbol{P}_4 &= \boldsymbol{A}_{22}(\boldsymbol{B}_{21} - \boldsymbol{B}_{11}),\\
\boldsymbol{P}_5 &= (\boldsymbol{A}_{11} + \boldsymbol{A}_{12})\boldsymbol{B}_{22},\\
\boldsymbol{P}_6 &= (\boldsymbol{A}_{21} - \boldsymbol{A}_{11})(\boldsymbol{B}_{11} + \boldsymbol{B}_{12}),\\
\boldsymbol{P}_7 &= (\boldsymbol{A}_{12} - \boldsymbol{A}_{22})(\boldsymbol{B}_{21} + \boldsymbol{B}_{22}),\\
\boldsymbol{C}_{11} &= \boldsymbol{P}_1 + \boldsymbol{P}_4 - \boldsymbol{P}_5 + \boldsymbol{P}_7,\\
\boldsymbol{C}_{12} &= \boldsymbol{P}_3 + \boldsymbol{P}_5,\\
\boldsymbol{C}_{21} &= \boldsymbol{P}_2 + \boldsymbol{P}_4,\\
\boldsymbol{C}_{22} &= \boldsymbol{P}_1 + \boldsymbol{P}_3 - \boldsymbol{P}_2 + \boldsymbol{P}_6.
\end{aligned}
$$

How many sub-matrix multiplications and sub-matrix additions are needed?

(c) Assume all sub-matrix multiplications and sub-matrix additions are calculated in a standard way. Since all sub-matrices are of size $\frac{n}{2} \times \frac{n}{2}$, we need $2\left(\frac{n}{2}\right)^3$ and $\left(\frac{n}{2}\right)^2$ flops for one sub-matrix multiplication and one sub-matrix addition respectively. How many flops are needed for the algorithms in Questions 4a and 4b respectively?

(d) (*Bonus Question.*) Suppose $n = 2^k$ for some integer k. We use the same algorithm as in Question 4b to each sub-matrix multiplications recursively until the sub-matrix size is reduced to 1×1. What is the computational complexity of the resulting matrix multiplication algorithm (known as *Stranssen matrix multiplication*)?

5. There are some other accelerations of matrix multiplications. A key idea in *Coppersmith–Winograd Matrix Multiplication* is to use the following trick to reduce the number multiplications. Suppose n is even and define

$$f(\boldsymbol{x}) = \sum_{i=1}^{n/2} x_{2i-1}x_{2i}, \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^n.$$

(a) Show that the inner product can be re-expressed as

$$\boldsymbol{x}^T\boldsymbol{y} = \sum_{i=1}^{n/2}(x_{2i-1} + y_{2i})(x_{2i} + y_{2i-1}) - f(\boldsymbol{x}) - f(\boldsymbol{y}), \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

(b) Now consider the matrix multiplication $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ for $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$. Give an algorithm for computing $\boldsymbol{C}$ that uses only $\frac{n^3}{2} + O(n^2)$ scalar multiplications. (*Notice that a standard matrix multiplication needs n^3 scalar multiplications. Together with other techniques, we can obtain Coppersmith–Winograd algorithm whose complexity is only $O(n^{2.375})$.*)

6. Use Gaussian Elimination (without row interchanges) to solve the following system of linear equations.

$$
\begin{cases}
2x_1 + x_2 + x_3 = 2\\
4x_1 + 5x_2 + 3x_3 = 2\\
2x_1 - 2x_2 + 3x_3 = 7.
\end{cases}
$$