Homework 1 solutions

COMS W3251, Summer Session 2

July 6, 2020

1 Matrix and vector computations (20 points)

For each of the following expressions, compute the result. You can use properties of basic operations to simplify the expression:

$$\left(\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix}^T\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix}+4I\right)e_2, \tag{1.1}$$

$$\left(\begin{bmatrix}2 & 1\\0 & -1\end{bmatrix}^T+\begin{bmatrix}0 & -1\\4 & 1\end{bmatrix}^T\right)^T\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix}, \tag{1.2}$$

$$\left\langle\left(6\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\\1 & 0 & 0\end{bmatrix}-3I\right)(2e_1+e_3),\begin{bmatrix}1\\-1\\1\end{bmatrix}\right\rangle, \tag{1.3}$$

$$\langle 2e_1+3e_2+4e_3, 8e_2-6e_3\rangle. \tag{1.4}$$

Solution:

1.

$$\left(\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix}^T\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix}+4I\right)e_2 =$$
$$\left(\begin{bmatrix}2 & 0\\3 & 1\\-1 & 1\end{bmatrix}\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix}+4I\right)e_2 =$$
$$\left(\begin{bmatrix}4 & 6 & -2\\6 & 10 & -2\\-2 & -2 & 2\end{bmatrix}+4I\right)e_2 =$$
$$\begin{bmatrix}8 & 6 & -2\\6 & 14 & -2\\-2 & -2 & 6\end{bmatrix}e_2 =$$
$$\begin{bmatrix}6\\14\\-2\end{bmatrix}$$

2.

$$\left(\begin{bmatrix}2 & 1\\0 & -1\end{bmatrix}^T + \begin{bmatrix}0 & -1\\4 & 1\end{bmatrix}^T\right)^T \begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix} =$$
$$\left(\begin{bmatrix}2 & 1\\0 & -1\end{bmatrix} + \begin{bmatrix}0 & -1\\4 & 1\end{bmatrix}\right) \begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix} =$$
$$\begin{bmatrix}2 & 0\\4 & 0\end{bmatrix} \begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\end{bmatrix} =$$
$$\begin{bmatrix}4 & 6 & -2\\8 & 12 & -4\end{bmatrix}$$

3.

$$\left\langle \left(6\begin{bmatrix}2 & 3 & -1\\0 & 1 & 1\\1 & 0 & 0\end{bmatrix} - 3I\right)(2e_1 + e_3), \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\rangle =$$
$$\left\langle \left(\begin{bmatrix}12 & 18 & -6\\0 & 6 & 6\\6 & 0 & 0\end{bmatrix} - 3I\right)\begin{bmatrix}2\\0\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\rangle =$$
$$\left\langle \begin{bmatrix}9 & 18 & -6\\0 & 3 & 6\\6 & 0 & -3\end{bmatrix}\begin{bmatrix}2\\0\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\rangle =$$
$$\langle \begin{bmatrix}12\\6\\9\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\rangle =$$
$$15$$

4.

$$\langle 2e_1 + 3e_2 + 4e_3, 8e_2 - 6e_3\rangle =$$
$$\langle \begin{bmatrix}2\\3\\4\end{bmatrix}, \begin{bmatrix}0\\8\\-6\end{bmatrix}\rangle =$$
$$0$$

2 Vector spaces (20 points)

Let $P_3(\mathbb{R})$ be the set of all polynomials of **degree 3 or less** with real coefficients. Prove that $P_3(\mathbb{R})$ is a vector space when considering $+$ as point-wise sum and $\cdot$ as point-wise scalar multiplication (that is, if $p_1, p_2 \in P_3(\mathbb{R})$, then $p_1 + p_2$ is the polynomial such that $p_1 + p_2(x) = p_1(x) + p_2(x)$, while $\alpha \cdot p_1$ is the polynomial such that $\alpha \cdot p_1(x) = \alpha p_1(x)$). Can you find a finite subset of $P_3(\mathbb{R})$ that spans $P_3(\mathbb{R})$?

Solution:

Properties of the sum:

1. **Commutativity:**
$$(p_1 + p_2)(x) = p_1(x) + p_2(x) = p_2(x) + p_1(x) = (p_2 + p_1)(x)$$

2. **Associativity:**
$$(p_1 + (p_2 + p_3))(x) = p_1(x) + p_2(x) + p_3(x) = ((p_1 + p_2) + p_3)(x)$$

3. **Identity element:** Let p_0 be the polynomial such that $p_0(x) = 0$ for every $x \in \mathbb{R}$. Then $p_0 \in P_3(\mathbb{R})$ and for every $p \in P_3(\mathbb{R})$ we have
$$(p + p_0)(x) = p(x) + p_0(x) = p(x)$$

4. **Existence of an inverse** For every $p \in P_3(\mathbb{R})$, let $-p$ be the polynomial such that $(-p)(x) = -p(x)$. Then
$$p + (-p)(x) = p(x) - p(x) = 0 = p_0(x)$$

Properties of the scalar multiplication:

1. **Compatibility with field multiplication:** Let $\alpha, \beta \in \mathbb{R}$. Then for every p
$$\alpha \cdot (\beta \cdot p)(x) = \alpha \cdot (\beta p(x)) = \alpha\beta p(x) = (\alpha\beta) \cdot p(x)$$

2. **Existence of identity element:**
$$1 \cdot p(x) = p(x)$$

3. **Distributivity with respect to +**: for every $\alpha \in \mathbb{R}$
$$\alpha \cdot (p_1 + p_2)(x) = \alpha p_1(x) + \alpha p_2(x) = (\alpha p_1 + \alpha p_2)(x)$$

4. **Distributivity with respect to scalar addition**: for every $\alpha, \beta \in \mathbb{R}$:
$$((\alpha + \beta) \cdot p)(x) = \alpha p(x) + \beta p(x) = (\alpha \cdot p)(x) + (\beta \cdot p)(x)$$

Finally, we need to prove that $P_3(\mathbb{R})$ is closed under linear combinations. That is, for every $\alpha, \beta \in \mathbb{R}$ and $p_1, p_2 \in P_3(\mathbb{R})$, then $\alpha p_1 + \beta p_2 \in P_3(\mathbb{R})$. But this is true because a linear combination of polynomials of degree less than or equal to 3 is still a polynomial of degree less than or equal to 3.

Finally, we know that all polynomials of degree less than or equal to 3 can be written as
$$p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3$$
for some $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ in $\mathbb{R}$. This means that every polynomial is a linear combination of 4 monomials $\{1, x, x^2, x^3\}$, so that $P_3(\mathbb{R}) = \text{span}(1, x, x^2, x^3)$.

3 Linear functions (25 points)

For each of the following function T, prove if they are linear or not. If they are, find the matrix A such that $T(x) = Ax$. If they are not, find a counterexample to the linearity condition.

$$T(x_1, x_2) = (x_1 + 4x_2, |x_2|) \tag{3.1}$$
$$T(x_1, x_2, x_3) = (x_1 - 8x_3, 0, 3x_2) \tag{3.2}$$
$$T(x_1, x_2, x_3) = (x_1 + 3, x_2, x_3 - x_1) \tag{3.3}$$
$$T(x_1, x_2) = x_1x_2 + x_2^2 \tag{3.4}$$
$$T(x_1) = (0, x_1, x_1^2, x_1^4) \tag{3.5}$$
$$T(x_1) = (x_1, x_1, 0, x_1) \tag{3.6}$$

Solution:

1. Not linear: absolute value in the second component is not linear. Consider $v_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then

$$T(v_1 + v_2) = T(\begin{bmatrix} 0 \\ 0 \end{bmatrix}) = (0, 0)$$

while

$$T(v_1) + T(v_2) = (-4, 1) + (4, 1) = (0, 2) \neq T(v_1 + v_2)$$

2. Linear. In fact we have

$$A = \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & -8 \\ 0 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

We can double-check by considering

$$\begin{bmatrix} 1 & 0 & -8 \\ 0 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - 8x_3 \\ 0 \\ 3x_2 \end{bmatrix}$$

3. Non-linear: $T(0, 0, 0) = (3, 0, 0) \neq (0, 0, 0)$. Therefore

$$T(2(0, 0, 0)) = T(0, 0, 0) = (3, 0, 0)$$

while

$$2T(0, 0, 0) = (6, 0, 0) \neq T(2(0, 0, 0))$$

4. Non-linear: the second component is quadratic. For $v_1 = (0, 1)$, we have

$$T(3v_1) = T(0, 3) = 9$$

while

$$3T(v_1) = 3 \neq T(3v_1)$$

5. Non-linear: the third and fourth component are respectively quadratic and a fourth power. For $v_1 = 1$, we have

$$T(2v_1) = T(2) = (0, 2, 4, 16)$$

while

$$2T(v_1) = 2T(1) = (0, 2, 2, 2)$$

6. Linear. In this case the matrix is given by

$$A = [T(e_1)] = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

4 Transformations of the plane (15 points)

Find the linear or affine transformations that satisfy the desired properties and write it in the form $T(x) = Ax + b$:

1. The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ sending the origin to itself and a triangle of vertices $(0, 0), (1, 0), (0, 1)$ to a triangle of vertices $(0, 0), (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$.
2. The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ sending the unit square $[0, 1] \times [0, 1]$ to the parallelogram with vertices $(4, 1), (4, 2), (7, 2), (7, 3)$.

Solutions:

1. This is a rotation of $\frac{\pi}{4}$ in the counter-clockwise direction. Therefore it is given by the matrix"

$$A = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

Notice that $A \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ since it's linear. Moreover

$$A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

$$A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

2. This is an affine transformation, since $T(0, 0) = (4, 1) \neq (0, 0)$. Therefore we know that $T(x) = Ax + \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. In order to find the residual matrix, we consider

a linear transformation $T_l(x) = T(x) - b$. Then $T_l(0,0) = (0,0)$, the origin is sent to the origin, $T_l(1,0) = T(1,0) - (4,1) = (7,2) - (4,1) = (3,1)$ and $T_l(0,1) = T(0,1) - (4,1) = (0,1)$. Then $T_l(x) = Ax$ for

$$A = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}$$

Therefore we have that

$$T(x) = T_l(x) + b = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

5 Flop count (20 points)

Let $A, B \in \mathbb{R}^{n \times n}$ and $x, y \in \mathbb{R}^n$. Following the order given by parenthesis, calculate the complexity of the following operations in term of flops.

$$\langle (A(Bx)), (\alpha y) \rangle \tag{5.1}$$

$$\alpha(\langle (AB)x, y \rangle) \tag{5.2}$$

Which one has the lowest complexity? Can you reduce the complexity even more?

Solutions:

1, non-sparse case: (Bx) is a matrix-vector product, so it has n inner products, each with $2n - 1$ flops.

(αy) is a scalar multiplication of a vector, so it has n flops.

$(A(Bx))$ is another matrix-vector product, so it has the same complexity as the previous one, with $n(2n - 1)$ flops.

Finally, we have the scalar product between two $n \times 1$ vectors, so there are $2n - 1$ flops.

$$tot = n(2n-1) + n + n(2n-1) + 2n - 1 = 4n^2 + n - 1 \sim 4n^2$$

2, non-sparse case: (AB) is a matrix-matrix product, so it has n^2 operations, each with $2n - 1$ flops.

$(AB)x$ is a matrix-vector product, so it has n operations, each with $2n - 1$ flops.

The inner product $\langle (AB)x, y \rangle$ has $2n - 1$ operations.

Finally, the multiplication by scalar α has 1 flop. Therefore

$$tot = n^2(2n-1) + n(2n-1) + 2n - 1 + 1 = 2n^3 + n^2 + n \sim 2n^3$$

Therefore the first method has a smaller computational complexity. Since there is always a matrix-vector product, the complexity cannot be lower than $n(2n - 1)$. However, a marginal improvement can be obtained by multipliying by α after taking the inner product.