Full name.: William Mardick-Kanter
Student id.: ....

1. Question:

Determine which of the following statements are **true** or **false**. Put a cross on the corresponding check box. Note that each correct answer is 1 point, each wrong answer is -1 point, and each unanswered one is 0 point. The total minimum number of points for this question is zero. *15 points*

true	false	
X	☐	Given two n-vectors a and b, they are orthogonal if $a^T b = 0$.
☐	X	Given an affine function $f(x)$, we have $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ for all scalar values of α, β and all n-vectors x, y.
☐	X	The function $f(x) = x_1 - x_2 + \ldots + (-1)^{n+1} x_n$ is a linear function, where $x = (x_1, \ldots, x_n)$.
☐	X	The first order Taylor approximation of $f(x)$ near point z is given by $\hat{f}(x) = f(z) + \nabla f(z)^T (x + z)$.
X	☐	Given $m \times n$ matrices A, B and n-vector x, the following inequality holds: $\lVert (A + B)x \rVert \leq \lVert Ax \rVert + \lVert Bx \rVert$.
X	☐	Performing Gram Schmidt orthogonalization for n vectors of size n takes around $2n^3$ flops.
X	☐	If one runs the k-means algorithm for minimizing the clustering objective J^{clust} for a sufficiently large number of iterations, it will eventually find the smallest possible value of J^{clust}.
☐	☐	The Gram Schmidt algorithm gives us n orthonormal vectors, if the input to the algorithm consists of n linearly independent vectors.
X	☐	$(A^T)^T = A$.
☐	X	Given an $m \times n$ matrix A and an $n \times 1$ vector v, their multiplication costs $2m^2 n$ flops.
☐	☐	Given an n-vector x, no more than 25% of entries can satisfy $\lvert x_i \rvert \geq 2\mathbf{rms}(x)$.
☐	X	The set $\{a_1, a_2, a_3\}$ is linearly independent, where a_1, a_2, a_3 are vectors of size 2 and all nonzero.
☐	X	The function $f(x_1, x_2, x_3, x_4) = (x_2, x_1, x_4, x_3)$ is linear.
☐	☐	For a given matrix A and vector b, equation $Ax = b$ always has a solution if A is wide.
X	☐	For a given vector b, the function $f(x) = b * x$ is linear, where $*$ denotes the convolution operator.

Problem 2 *worth 5 points*
Consider a directed graph $G(V,E)$ (or a relation on n objects, where the objects are members of set $V, |V| = n$, and the relation is E) represented as a matrix A. We define another Graph (relation) on the same set of objects, $G'(V,E')$, such that if $(x,y) \in E$, the $(x,y) \notin E'$ and if $(x,y) \notin E$ then $(x,y) \in E'$, represent it as a matrix A'. Express A' in terms of A.

Solution:

$$A' = \mathbf{1} - A$$

Let's use a simple 3×3 matrix that represents $G(V,E)$. $G(V,E) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. If we have a matrix of relations that are expressly not in G we would see a matrix that looks like $G'(V,E') = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. Let's plug this into $A' = |\mathbf{1} - A|$ with $A' = G'$ and $A = G$.

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \mathbf{1} - \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1-1 & 1-0 & 1-1 \\ 1-0 & 1-1 & 1-0 \\ 1-1 & 1-0 & 1-1 \end{bmatrix}$$

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

Problem 3 *worth 10 points*

Apply the Gram-Schmidt algorithm to vectors $a_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $a_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$, and $a_3 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$ and *provide the corresponding orthonormal set of vectors.*

Solution:

$$\tilde{q}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$\tilde{q}_1 = \frac{\tilde{q}_1}{||\tilde{q}_1||}$$
$$\tilde{q}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\tilde{q}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - (\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}^T \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}) \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$
$$\tilde{q}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - (\frac{-1}{\sqrt{3}}) \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$
$$\tilde{q}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$
$$\tilde{q}_2 = \begin{bmatrix} \frac{-6}{3} + \frac{1}{3} \\ 0 + \frac{1}{3} \\ \frac{3}{3} + \frac{1}{3} \end{bmatrix}$$
$$\tilde{q}_2 = \begin{bmatrix} \frac{-5}{3} \\ \frac{1}{3} \\ \frac{4}{3} \end{bmatrix} \neq 0$$
$$\tilde{q}_2 = \begin{bmatrix} \frac{-5}{3} \\ \frac{1}{3} \\ \frac{4}{3} \end{bmatrix} /||\tilde{q}_2||$$
$$||\tilde{q}_2|| = (\frac{-5}{3})^2 + (\frac{1}{3})^2 + (\frac{4}{3})^2$$
$$||\tilde{q}_2|| = \sqrt{\frac{42}{9}} = \frac{\sqrt{42}}{3}$$
$$\tilde{q}_2 = \frac{1}{\sqrt{42}} \begin{bmatrix} -5 \\ 1 \\ 4 \end{bmatrix}$$

$$\tilde{q}_3 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} - (\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}^T \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}) \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} - (\frac{1}{\sqrt{42}} \begin{bmatrix} -5 \\ 1 \\ 4 \end{bmatrix}^T \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}) \frac{1}{\sqrt{42}} \begin{bmatrix} -5 \\ 1 \\ 4 \end{bmatrix}$$

$$\tilde{q}_3 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} - (\sqrt{3}) \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} - (\frac{3}{\sqrt{42}}) \begin{bmatrix} \frac{-5}{\sqrt{42}} \\ \frac{1}{\sqrt{42}} \\ \frac{4}{\sqrt{42}} \end{bmatrix}$$

$$\tilde{q}_3 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} - (\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{-15}{42} \\ \frac{3}{42} \\ \frac{12}{42} \end{bmatrix}$$

$$\tilde{q}_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} - \begin{bmatrix} \frac{-15}{42} \\ \frac{3}{42} \\ \frac{12}{42} \end{bmatrix}$$

$$\tilde{q}_3 = \begin{bmatrix} \frac{-42}{42} \\ \frac{84}{42} \\ \frac{-42}{42} \end{bmatrix} - \begin{bmatrix} \frac{-15}{42} \\ \frac{3}{42} \\ \frac{12}{42} \end{bmatrix}$$

$$\tilde{q}_3 = \begin{bmatrix} \frac{-27}{42} \\ \frac{81}{42} \\ \frac{-54}{42} \end{bmatrix} \neq 0$$

$$\tilde{q}_3 = \begin{bmatrix} \frac{-27}{42} \\ \frac{81}{42} \\ \frac{-54}{42} \end{bmatrix} /||\tilde{q}_3||$$

Problem 4 *worth 10 points*

Find the Taylor approximation $\hat{f}(x_1, x_2, x_3)$ of $f(x_1, x_2, x_3) = \begin{bmatrix} x_1x_2 \\ x_2x_3 \\ x_3x_1 \end{bmatrix}$ at $z = (1, 1, 1)$. Then, find the value of $\hat{f}(0, 1, 1)$.

Solution: The taylor approximation for this function would look like:

$$\hat{f}(x)_1 = f_1(z) + \frac{\delta f_1(z)}{\delta x_1}(x_1 - z_1) + \frac{\delta f_1(z)}{\delta x_2}(x_2 - z_2) + \frac{\delta f_1(z)}{\delta x_3}(x_3 - z_3)$$
$$\hat{f}(x)_2 = f_2(z) + \frac{\delta f_2(z)}{\delta x_1}(x_1 - z_1) + \frac{\delta f_2(z)}{\delta x_2}(x_2 - z_2) + \frac{\delta f_2(z)}{\delta x_3}(x_3 - z_3)$$
$$\hat{f}(x)_3 = f_3(z) + \frac{\delta f_3(z)}{\delta x_1}(x_1 - z_1) + \frac{\delta f_3(z)}{\delta x_2}(x_2 - z_2) + \frac{\delta f_3(z)}{\delta x_3}(x_3 - z_3)$$

For a taylor series on a vector the $f(x)_i$ refers to the i^{th} element in the function's vector.

$$\hat{f}(x)_1 = f_1(z) + \frac{\delta f_1(z)}{\delta x_1}(x_1 - z_1) + \frac{\delta f_1(z)}{\delta x_2}(x_2 - z_2) + \frac{\delta f_1(z)}{\delta x_3}(x_3 - z_3)$$
$$\hat{f}(x)_1 = 1 + x_2(0 - 1) + x_1(1 - 1) + 0(1 - 1)$$
$$\hat{f}(x)_1 = 1 + 1(-1) + 0 + 0$$
$$\hat{f}(x)_1 = 1 - 1$$

So the first element of $\hat{f}(x)$ is 0.

$$\hat{f}(x)_2 = f_2(z) + \frac{\delta f_2(z)}{\delta x_1}(x_1 - z_1) + \frac{\delta f_2(z)}{\delta x_2}(x_2 - z_2) + \frac{\delta f_2(z)}{\delta x_3}(x_3 - z_3)$$
$$\hat{f}(x)_2 = 1 + 0(0 - 1) + x_3(1 - 1) + x_2(1 - 1)$$
$$\hat{f}(x)_2 = 1 + 0 + 0 + 0$$

The second element of $\hat{f}(x)$ is 1.

$$\hat{f}(x)_3 = f_3(z) + \frac{\delta f_3(z)}{\delta x_1}(x_1 - z_1) + \frac{\delta f_3(z)}{\delta x_2}(x_2 - z_2) + \frac{\delta f_3(z)}{\delta x_3}(x_3 - z_3)$$
$$\hat{f}(x)_3 = 1 + x_3(0 - 1) + 0(1 - 1) + x_1(1 - 1)$$
$$\hat{f}(x)_3 = 1 + 1(0 - 1) + 0(1 - 1) + 0(1 - 1)$$
$$\hat{f}(x)_3 = 1 - 1$$

And the third element of $\hat{f}(x)$ is 0. So according to the taylor approximation our function should produce a vector that looks like $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Now evaluating the vector $\begin{bmatrix}0\\1\\1\end{bmatrix}$ in our function $f(x_1, x_2, x_3) = \begin{bmatrix}x_1x_2\\x_2x_3\\x_3x_1\end{bmatrix}$ we get

$$f(x_1, x_2, x_3) = \begin{bmatrix}x_1x_2\\x_2x_3\\x_3x_1\end{bmatrix}$$
$$f(0,1,1) = \begin{bmatrix}0\times 1\\1\times 1\\1\times 0\end{bmatrix}$$
$$f(0,1,1) = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

This accuracy is because of how close the vectors are to each other.

Problem 5 *worth 10 points*
We define the determinate of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ as $\mathbf{det}(A) = ad - bc$. For example for $A = \begin{bmatrix} 3 & 10 \\ 2 & 15 \end{bmatrix}$, $\mathbf{det}(A) = 3(15) - 10(2) = 45 - 20 = 25$. Now consider a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Show that if the rows of A are linearly dependent, then the columns of A are also linearly dependent as well. Show that if the rows are linearly dependent then $\mathbf{det}(A) = 0$.

Solution: First we break the matrix into four vectors:

$$c_1 = \begin{bmatrix} a \\ c \end{bmatrix}, \quad c_2 = \begin{bmatrix} b \\ d \end{bmatrix}, \quad r_1 = \begin{bmatrix} a \\ b \end{bmatrix}, \quad r_2 = \begin{bmatrix} c \\ d \end{bmatrix}$$

If we have a collection of linearly dependent vectors then we could multiply a scalar by these vectors and get another vector from the set as our product. Also, as given, the determinate of this matrix would be $ad - bc = 0$ which means $ad = bc$. Rearranging this we get $\frac{a}{b} = \frac{c}{d}$. If the columns are linearly dependent then there exists a scalar such that $\beta c_2 = c_1$. let's let $\beta = \frac{a}{b}$.

$$\beta \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$$
$$\begin{bmatrix} \frac{a}{b} \times b \\ \frac{a}{b} \times d \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$$
$$\begin{bmatrix} a \\ \frac{ad}{b} \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$$

Going back to $ad - bc = 0$, we see that c is equal to $\frac{ad}{b}$. So we have shown $\beta c_2 = c_1$.
Using this same principle we show our rows are linearly dependent as well, this time $\gamma = \frac{d}{b}$ to prove $\gamma r_1 = r_2$

$$\gamma \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$
$$\begin{bmatrix} \frac{d}{b} \times a \\ \frac{d}{b} \times b \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$
$$\begin{bmatrix} \frac{ad}{b} \\ d \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$

As shown earlier, $c = \frac{ad}{b}$, so we have multiplied a scalar by our row, and got the other row in return.
To show $\mathbf{det}(A) = 0$ I will make the substitution be $a = \frac{bc}{d}$

$$ad - bc = 0$$
$$\frac{bc}{d}d - bc = 0$$
$$bc - bc = 0$$

Problem 6 *worth 5 points*
Assume $A = \begin{bmatrix} 305 & 304 \\ 238 & 250 \end{bmatrix}$, $x = \begin{bmatrix} 2020 \\ -311 \end{bmatrix}$, amd $y = \begin{bmatrix} -2019 \\ 310 \end{bmatrix}$. Find $Ax + Ay$

Solution:

$$\begin{aligned} Ax + Ay \\ A(x+y) \\ A(\begin{bmatrix} 2020 \\ -311 \end{bmatrix} + \begin{bmatrix} -2019 \\ 310 \end{bmatrix}) \\ A(\begin{bmatrix} 1 \\ -1 \end{bmatrix}) &= \begin{bmatrix} 305 & 304 \\ 238 & 250 \end{bmatrix} (\begin{bmatrix} 1 \\ -1 \end{bmatrix}) \\ &= \begin{bmatrix} 305(1) + 304(-1) \\ 238(1) + 250(-1) \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ -12 \end{bmatrix} \end{aligned}$$

Problem 7 *worth 5 points*
Consider an $n \times n$ matrix A and $n \times 1$ vectors v and u. Approximate the complexity (in terms of number of flops) taken to compute $u^T(Av)$.

Solution: Complexity of matrix vector multiplication: $2mn$, where m is number of rows and n is number of columns, here our rows and columns are equal so we get $2n^2$
Complexity of transposing two n-vectors (which Ax will produce): $2n + 1$
Summed: $2n^2 + 2n$, drop the 1 because computers.