1. Here is another way to derive the normal equation without using any calculus. Recall that the null space or kernel of a matrix \( A \in \mathbb{R}^{m \times n} \) is the set 
\[
\ker(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}
\]
while the range space or image is the set 
\[
\text{im}(A) = \{ y \in \mathbb{R}^m \mid y = Ax \text{ for some } x \in \mathbb{R}^n \}
\]
and \( b \in \mathbb{R}^m \).
(a) Show that 
\[
\ker(A^\top A) = \ker(A).
\]
(b) Show that 
\[
\text{im}(A^\top A) = \text{im}(A^\top).
\]
(c) Deduce that 
\[
A^\top Ax = A^\top b
\]
always has a solution. We call this the normal equation.
(d) Give an example where \( Ax = b \) has no solution but \( A^\top Ax = A^\top b \) has a solution.
(e) Show that (a), (b), and (c) are false in general over a field with two elements \( \mathbb{F}_2 = \{0, 1\} \) with arithmetic done modulo 2.

2. We would like to solve the differential equation
\[
\begin{cases}
-v''(x) = \frac{m\omega^2}{k} v(x), & 0 < x < 1, \\
v(0) = 0, v(1) = 0.
\end{cases}
\]
This comes up when studying a vibrating string with \( m \) the mass per unit length and \( k \) the stiffness per unit length, both positive constants. We need to determine the function \( v : [0, 1] \to \mathbb{R} \) and the number \( \omega \in \mathbb{R} \). Here \( v(x) \) gives us the amplitude of the string at \( x \) and and \( \omega \) gives us the vibration frequency of the string.
(a) Following the technique used in Lecture 3, show that we may discretize the differential equation into the following difference equation
\[
\begin{cases}
\frac{-v_{i-1} + 2v_i - v_{i+1}}{n^{-2}} = \lambda v_i, & 1 \leq i \leq n-1, \\
v_0 = 0, v_n = 0.
\end{cases}
\]
(b) Show that the difference equation can be rewritten as an eigenvalue problem
\[
Av = \lambda v
\]
where \( \lambda \) is an approximation of \( m\omega^2/k \).
3. Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$. Suppose we would like to learn a function $f : X \rightarrow Y$ from a training set of data $\{(x_i, y_i) \in X \times Y \mid i = 1, \ldots, m\}$. We will assume that $f$ can be expressed as a linear combination

$$f(x) = \sum_{i=1}^{m} c_i K(x, x_i)$$

where $K(x, y) = \exp(-\|x - y\|^2/2\sigma^2)$ is a Gaussian kernel. Following the data fitting technique discussed in Lecture 3, describe how one may determine the coefficients $c_1, \ldots, c_m \in \mathbb{R}$ by solving a least squares problem. You will need to describe the least squares problem explicitly: What are the coefficient matrix and the right-hand side.

4. In testing your codes, it is often important to know how to randomly generate matrices with some specified properties. In MATLAB, you can generate a random $m \times n$ matrix $X$ with built-in functions `rand(m,n)` and `randn(m,n)`, where the entries are drawn respectively from the uniform distribution on the interval $(0, 1)$ and the standard normal distribution. For each of the following, write a program that will generate:

(a) $n \times n$ real symmetric matrices, i.e. $X^\top = X$;
(b) $n \times n$ real skew-symmetric matrix, i.e. $X^\top = -X$;
(c) $n \times n$ non-singular matrices (a.k.a. invertible matrices);
(d) $n \times n$ symmetric positive definite Toeplitz matrices;
(e) $m \times n$ matrices of rank $r$, where $r \in \{0, 1, \ldots, \min(m, n)\}$ is an unspecified input;
(f) $m \times n$ matrices whose entries are uniformly distributed in $[\alpha, \beta]$, where $\alpha < \beta$ are unspecified inputs;
(g) $m \times n$ matrices whose entries are normally distributed with mean $\mu$ and variance $\sigma^2$, where $\mu$ and $\sigma$ are unspecified inputs;
(h) $m \times n$ matrices whose entries are either 0 or 1 with probabilities $p$ and $1 - p$ respectively, where $p \in (0, 1)$ is an unspecified input.