Math Foundations of ML, Fall 2018

Homework #3

Due Friday September 14, at the beginning of class

As stated in the syllabus, unauthorized use of previous semester course materials is strictly prohibited in this course.

1. Using you class notes, prepare a 1-2 paragraph summary of what we talked about in class in the last week. I do not want just a bulleted list of topics, I want you to use complete sentences and establish context (Why is what we have learned relevant? How does it connect with other things you have learned here or in other classes?). The more insight you give, the better.

2. A square $N \times N$ matrix $\boldsymbol{G}$ is *invertible* if
$$\boldsymbol{x}_1 \neq \boldsymbol{x}_2 \;\Leftrightarrow\; \boldsymbol{G}\boldsymbol{x}_1 \neq \boldsymbol{G}\boldsymbol{x}_2.$$
That is, $\boldsymbol{G}\boldsymbol{x}$ is different for every different $\boldsymbol{x}$.

   (a) It is clear that $\boldsymbol{G}\boldsymbol{0} = \boldsymbol{0}$ for any matrix $\boldsymbol{G}$. Show that if
$$\boldsymbol{G}\boldsymbol{\alpha} = \boldsymbol{0} \;\Rightarrow\; \boldsymbol{\alpha} = \boldsymbol{0},$$
then $\boldsymbol{G}$ must be invertible.

   (b) Let $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N$ be N linearly independent vectors in a Hilbert space, and let $\mathcal{T} = \operatorname{span}\{\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N\}$. Show that if $\boldsymbol{z} \in \mathcal{T}$ and $\langle \boldsymbol{z}, \boldsymbol{\psi}_n \rangle = 0$ for all $n = 1, \ldots, N$, then it must be true that $\boldsymbol{z} = \boldsymbol{0}$. (Hint: take the inner product of $\boldsymbol{z}$ with itself.)

   (c) Show that if $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N$ are N linearly independent vectors in a Hilbert space, then the Gram matrix
$$\boldsymbol{G} = \begin{bmatrix} \langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_1 \rangle & \langle \boldsymbol{\psi}_2, \boldsymbol{\psi}_1 \rangle & \cdots & \langle \boldsymbol{\psi}_N, \boldsymbol{\psi}_1 \rangle \\ \langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_2 \rangle & \langle \boldsymbol{\psi}_2, \boldsymbol{\psi}_2 \rangle & & \langle \boldsymbol{\psi}_N, \boldsymbol{\psi}_2 \rangle \\ \vdots & & \ddots & \vdots \\ \langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_N \rangle & \cdots & & \langle \boldsymbol{\psi}_N, \boldsymbol{\psi}_N \rangle \end{bmatrix},$$
is invertible.

3. In this problem, we will develop the computational framework for approximating a continuous-time signal on $[0, 1]$ using scaled and shifted version of the classic bell-curve bump:
$$\phi(t) = e^{-t^2}.$$
Fix an integer $N > 0$ and define $\phi_k(t)$ as
$$\phi_k(t) = \phi\left(\frac{t - (k - 1/2)/N}{1/N}\right) = \phi\left(Nt - k + 1/2\right)$$
for $k = 1, 2, \ldots, N$. The $\{\phi_k(t)\}$ are a basis for the subspace
$$\mathcal{T}_N = \operatorname{span}\{\boldsymbol{\phi}_k\}_{k=1}^{N}.$$

(a) For a fixed value of N, we can plot all of the $\phi_k(t)$ on the same set of axes in MATLAB using:

```
phi = @(z) exp(-z.^2);
t = linspace(0, 1, 1000);
figure(1); clf
hold on
for kk = 1:N
   plot(t, phi(N*t - kk + 1/2))
end
```

OR in Python using:

```
import numpy as np
import matplotlib.pyplot as plt

phi = lambda z: np.exp(-z**2)
t = np.linspace(0,1,1000)

plt.figure(1)
plt.clf()
for kk in range(N):
    plt.plot(t, phi(N*t - kk + 0.5))
```

Do this for $N = 10$ and $N = 25$ and turn in your plots.

(b) Since $\{\boldsymbol{\phi}_k\}$ is a basis for $\mathcal{T}_N$, we can write any $\boldsymbol{y} \in \mathcal{T}_N$ as

$$y(t) = \sum_{k=1}^{N} a_k \phi_k(t)$$

for some set of coefficients $a_1, \ldots, a_N \in \mathbb{R}^N$. If these coefficients are stacked in an N-vector `a` in MATLAB, we can plot $y(t)$ using

```
t = linspace(0,1,1000);
y = zeros(size(t));
for jj = 1:N
   y = y + a(jj)*phi(N*t - jj + 1/2);
end
plot(t, y)
```

OR in python, we can plot $y(t)$ using

```
y = np.zeros(1000)

for jj in range(N):
    y = y+a[jj]*phi(N*t - jj + 0.5)

plt.figure()
plt.plot(t,y)
```

Do this for $N = 4$, and $a_1 = -1, a_2 = 1, a_3 = 2, a_4 = -1/2$ and turn in your plot.

(c) Define the continuous-time signal $x(t)$ on $[0, 1]$ as

$$x(t) = \begin{cases} 4t & 0 \le t < 1/4 \\ -4t + 2 & 1/4 \le t < 1/2 \\ -\sin(20\pi t) & 1/2 \le t \le 1 \end{cases}.$$

Write MATLAB code that finds the closest point $\hat{\boldsymbol{x}}$ in $\mathcal{T}_N$ to $\boldsymbol{x}$ for any fixed N. By "closest point", we mean that $\hat{x}(t)$ is the solution to

$$\min_{\boldsymbol{y} \in \mathcal{T}_N} \|x(t) - y(t)\|_{L_2([0,1])}.$$

Turn in your code and four plots; one of which has $x(t)$ and $\hat{x}(t)$ plotted on the same set of axes for $N = 5$, and then repeat for $N = 10, 20$, and 50.
Hint: You can create a function pointer for $x(t)$ in MATLAB using

```
x = @(z) (z < 1/4).*(4*z) + (z>=1/4).*(z<1/2).*(-4*z+2) - (z>=1/2).*sin(20*pi*z);
```

OR in python using

```
x = lambda z: (z < .25)*(4*z) + (z >= 0.25)*(z < 0.5)*(-4*z+2) - \
 (z>= 0.5)*np.sin(20*np.pi*z)
```

and then calculate the continuous-time inner product $\langle \boldsymbol{x}, \boldsymbol{\phi}_k \rangle$ in MATLAB with

```
x_phik = @(z) x(z).*phi(N*z - jj + 1/2);
integral(x_phik, 0, 1)
```

OR in Python with

```
import scipy.integrate as integrate
x_phik = lambda z: x(z)*phi(N*z - jj + 0.5)
integrate.quad(x_phik, 0, 1)
```

You can use similar code to calculate the entries of the Gram matrix $\langle \boldsymbol{\phi}_j, \boldsymbol{\phi}_k \rangle$. (There is actually a not-that-hard way to calculate the $\langle \boldsymbol{\phi}_j, \boldsymbol{\phi}_k \rangle$ analytically that you can derive if you are feeling industrious — just think about what happens when you convolve a bump with itself.)

4. Do the exercise on page 55 of the notes (part 2 of approximating e^t with a quadratic polynomial).

5. In this question, we explore an algorithm that takes a basis for a subspace and produces and orthobasis for that same subspace.

(a) Let $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N$ be vectors in a Hilbert space with $\|\boldsymbol{\psi}_n\| > 0$. Show that if $\langle \boldsymbol{\psi}_i, \boldsymbol{\psi}_n \rangle = 0$ for all $i \neq n$, then these vectors a lineraly independent.

(b) Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ be a basis for a subspace $\mathcal{T}$ of a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ with induced norm $\|\cdot\|$. Define

$$\boldsymbol{\psi}_1 = \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|},$$

then for $k = 2, \ldots, N$,

$$\begin{aligned} \boldsymbol{u}_k &= \boldsymbol{v}_k - \sum_{\ell=1}^{k-1} \langle \boldsymbol{v}_k, \boldsymbol{\psi}_\ell \rangle \boldsymbol{\psi}_\ell, \\ \boldsymbol{\psi}_k &= \frac{\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|}. \end{aligned}$$

Argue that $\operatorname{span}(\{\boldsymbol{\psi}_k\}_{k=1}^N) \subset \mathcal{T}$. Now argue that $\langle \boldsymbol{\psi}_m, \boldsymbol{\psi}_n \rangle = 1$ if $m = n$ and is 0 if $m \neq n$. Combine these arguments into a proof that $\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_N$ is an orthobasis for $\mathcal{T}$.