AMATH 563 - Homework 1 (Theory)

Nate Whybra

May 2025

Problem 1 (a)

By the representer theorem, we know that for some $\alpha \in \mathbb{R}^N$:

$$f^*(x) = K(x, X)\alpha$$

Therefore:

$$\|f^*\|_H^2 = \langle f^*, f^* \rangle_H = \alpha^T K(X, X)\alpha$$

Also, for every x_j:

$$f^*(x_j) = K(x_j, X)\alpha$$

Now define $z_j^* = f^*(x_j) = K(x_j, X)\alpha$, then for $z^* \in \mathbb{R}^N$, we see:

$$z^* = K(X, X)\alpha \implies \alpha = K(X, X)^{-1} z^*$$

So that:

$$f^*(x) = K(x, X)K(X, X)^{-1} z^*$$

$$\|f\|_H^2 = (z^*)^T K(X, X)^{-1} K(X, X) K(X, X)^{-1} z^* = (z^*)^T K(X, X)^{-1} z^*$$

And:

$$z^* = \operatorname{argmin}_{z \in \mathbb{R}^n} \left(L(z_1, \cdots, z_n) + \frac{\lambda}{2} z^T K(X, X)^{-1} z \right)$$

Which immediately follows from the definition of our construction for z^*, and the formula for $\|f\|_H^2$ above.

Problem 1 (b)

Define with e_i being the $i-th$ standard basis vector in $\mathbb{R}^N$:

$$\psi_i(x) = K(x, X)K(X, X)^{-1}e_i$$

Then:

$$\begin{aligned}
f^*(x) &= K(x, X)K(X, X)^{-1}z^* \\
&= \sum_{i=1}^{n} K(x, X)K(X, X)^{-1}z_i^* e_i \\
&= \sum_{i=1}^{n} z_i^* K(x, X)K(X, X)^{-1}e_i \\
&= \sum_{i=1}^{n} z_i^* \psi_i(x)
\end{aligned}$$

This is desired. Now all we must show is that the $\psi_i's$ we defined satisfy the necessary properties. Firstly as $K(X, X)^{-1}$ is the inverse:

$$\psi_i(x_j) = K(x_j, X)K(X, X)^{-1}e_i = e_j^T e_i = \delta_{i,j}(x)$$

Lastly, we need to argue:

$$\psi_i = \arg\min_{\psi \in \mathcal{H}} \|\psi\|_{\mathcal{H}}^2 \quad \text{subject to } \psi(x_j) = \delta_{ij} \text{ for } j \in \{1, \ldots, N\}$$

To justify this, we observe the following. By the representer theorem again, the minimizer of any such constrained problem over $\mathcal{H}$ lies in the finite-dimensional subspace:

$$\mathcal{H}_N := \text{span}\{K(x, x_1), \ldots, K(x, x_N)\}$$

So we write:

$$\psi(x) = K(x, X)\alpha$$

And the "interpolation conditions" $\psi(x_j) = \delta_{ij}$ for all j become:

$$K(X, X)\alpha = e_i$$

So as $K(X, X)$ is invertible, this system has the unique solution:

$$\alpha = K(X, X)^{-1}e_i$$

And we see (just like above):

$$\psi_i(x) = K(x, X)K(X, X)^{-1}e_i$$

To confirm this is the minimum norm solution, let $\tilde{\psi} \in \mathcal{H}$ be any other function satisfying the same interpolation constraints. Write $\tilde{\psi} = \psi_i + \eta$ where $\eta \in \mathcal{H}$ and $\eta(x_j) = 0$ for all j. Since $\eta \perp \mathcal{H}_N$ and $\psi_i \in \mathcal{H}_N$, we have by Pythagoras:

$$\|\tilde{\psi}\|_{\mathcal{H}}^2 = \|\psi_i\|_{\mathcal{H}}^2 + \|\eta\|_{\mathcal{H}}^2 \geq \|\psi_i\|_{\mathcal{H}}^2$$

Thus ψ_i is the unique minimizer and we are done.

Problem 1 (c)

Since H is a RKHS, the reproducing property allows us to say that, as $\psi_i \in H$, that:

$$\langle \psi_i(x), K(x, x_j) \rangle_H = \psi_i(x_j) = \delta_{ij}$$

Therefore ψ_i and $K(x, x_j)$ are orthogonal whenever $i \neq j$ implying $\psi_i(x) \perp \text{span}\{K(x, x_j)\}_{i \neq j}$ as desired.

Problem 2 (a)

First suppose (λ, v) is an eigenpair of $\hat{L}$. Then:

$$\hat{L}v = \lambda v \quad \text{and} \quad Lu = \lambda Du$$

So then:

$$D^{-0.5}Lu = \lambda D^{0.5}Du = \lambda D^{-0.5}u$$

$$\implies D^{-0.5}LD^{-0.5}D^{0.5}u = \lambda D^{0.5}D^{-0.5}D^{0.5}u \quad (\text{inserting } I = D^{-0.5}D^{0.5})$$

$$\implies \hat{L}\left(D^{0.5}u\right) = \lambda(D^{0.5}u)$$

So (λ, v) with $v = D^{0.5}u$ or $u = D^{-0.5}v$ is an eigenpair for $\hat{L}$. Conversely, suppose $v = D^{0.5}u$, where $Lu = \lambda Du$, then $u = D^{-0.5}v$ and:

$$\begin{aligned} Lu &= \lambda Du \\ L(D^{-0.5}v) &= \lambda D\left(D^{-0.5}v\right) \\ D^{-0.5}LD^{0.5}v &= D^{-0.5}\lambda D^{0.5}v \\ \hat{L}v &= \lambda v \end{aligned}$$

So that (λ, v) is an eigenpair for $\hat{L}$, and we are done.

Problem 2 (b)

As G has M connected components, from lecture, the Normalized Graph Laplacian $\hat{L}$ can be written as diag$([\hat{L}_1, \cdots, \hat{L}_M])$ (a diagonal block of sub-Laplacian matrices for each connected component in G) and the Un-Normalized Graph Laplacian can L can be written as diag$([L_1, \cdots, L_M])$. Using Theorem 20.1 from Lecture, we have that $dim(null(L)) = M$ with $null(L) = \text{span}(\{1_{G_1}, \cdots, I_{G_M}\})$, where 1_{G_i} is the indicator function for the connected components G_i of G.

Now suppose $u \in null(L)$ so that from linearity $v = D^{0.5}u \in D^{0.5}null(L)$, then:

$$\begin{aligned}\hat{L}v &= D^{-0.5}LD^{-0.5}D^{0.5}u \\ &= D^{-0.5}Lu \\ &= 0\end{aligned}$$

So $v \in null(\hat{L})$ and $D^{0.5}null(L) \subset null(\hat{L})$. Conversely, suppose $v \in null(\hat{L})$, then:

$$\begin{aligned}\hat{L}v &= 0 \\ D^{0.5}\left(D^{-0.5}LD^{-0.5}v\right) &= 0 \\ LD^{-0.5}v &= 0 \\ Lu &= 0\end{aligned}$$

Therefore $u = D^{-0.5}v \in null(L)$ and $v = D^{0.5}u \in D^{0.5}null(L)$ so that $null(\hat{L}) \subset D^{0.5}null(L)$. Since the subsets go both ways, we have $null(\hat{L}) = D^{0.5}null(L)$. Now as $D^{0.5}$ is invertible, the null spaces of L and $\hat{L}$ both have dimension M, and finally we can say as desired:

$$null(\hat{L}) = \text{span}\left(\{D^{0.5}1_{G_1}, \cdots, D^{0.5}1_{G_M}\}\right)$$

Problem 2 (c)

We want to show that the eigenvalues of the normalized Laplacian $\hat{L} = D^{-1/2} L D^{-1/2}$ are bounded above by 2. That is, for all $j \leq N$:

$$\lambda_j \leq 2$$

The eigenvalues of a symmetric matrix like $\hat{L}$ can be written using the Courant–Fischer–W theorem. In particular, the largest eigenvalue satisfies:

$$\lambda_N = \max_{x \in \mathbb{R}^N \setminus \{0\}} \frac{x^T \hat{L} x}{x^T x}$$

Now since $\hat{L} = I - D^{-1/2} W D^{-1/2}$, we have:

$$\frac{x^T \hat{L} x}{x^T x} = 1 - \frac{x^T D^{-1/2} W D^{-1/2} x}{x^T x}$$

Letting $y = D^{-1/2} x$, the expression becomes:

$$\frac{x^T \hat{L} x}{x^T x} = 1 - \frac{y^T W y}{y^T D y}$$

To get an upper bound, we show that $\frac{y^T W y}{y^T D y} \geq -1$. Since W is symmetric with non-negative entries:

$$\begin{aligned}
y^T W y &= \sum_{i,j} w_{ij} y_i y_j \\
&\geq -\frac{1}{2} \sum_{i,j} w_{ij} (y_i^2 + y_j^2) \\
&= -\sum_i y_i^2 \sum_j w_{ij} = -\sum_i d_i y_i^2 = -y^T D y
\end{aligned}$$

So:

$$\frac{y^T W y}{y^T D y} \geq -1 \quad \Rightarrow \quad \frac{x^T \hat{L} x}{x^T x} \leq 1 - (-1) = 2$$

This holds for all $x \neq 0$, so all eigenvalues satisfy $\lambda_j \leq 2$. Now we show that the same is not true for the Un-Normalized Laplacian $L = D - W$. The maximum eigenvalue is:

$$\lambda_N = \max_{x \in \mathbb{R}^N \setminus \{0\}} \frac{x^T L x}{x^T x} = \max_{x \neq 0} \left(\frac{x^T D x}{x^T x} - \frac{x^T W x}{x^T x} \right) = \frac{2 x^T D x}{x^T x} = 2 \frac{\sum_{i=1}^N D_{ii} x_i^2}{x^T x}$$

Since the degrees D_{ii} can grow arbitrarily large as $N \to \infty$, the eigenvalues of L cannot be uniformly bounded.