HUT , Institute of mathematics Gripenberg
Mat-1.196 Mathematics of neural networks
Exercise 8
5.3–15.3.2002

1. Let H be a separable Hilbert space, let $(f_n)_{n=1}^{\infty}$ be a frame H, and let W be a closed subspace of H. Construct a frame in W using $(f_n)_{n=1}^{\infty}$ and the orthogonal projection of H onto W.
Solution: Denote the orthogonal projection $H \to W$ by P. Now we have to show that $(Pf_n)_{n]1}^{\infty}$ is a frame in W. Let $f \in W$ be arbitrary. Since $(f_n)_{n=1}^{\infty}$ is a frame in H and $f \in H$ (because $W \subset H$) we have

$$A\|f\|^2 \le \sum_{n=1}^{\infty} |\langle f_n, f\rangle|^2 \le B\|f\|^2.$$

But now $f = Pf$ and so that $\langle f_n, f\rangle = \langle f_n, Pf\rangle = \langle Pf_n, f\rangle$, and it follows immediately that $(Pf_n)_{n]1}^{\infty}$ is a frame in W.

---

2. Why is it not possible to construct a frame in $L^2(\mathbb{R}^d)$ of functions of the form $\frac{1}{\sqrt{a}}\varphi(\frac{\mathbf{u}\cdot\mathbf{x}-b}{a})$ when $d > 1$.
Solution: Suppose that there are numbers $\delta > 0$ and $c \in \mathbb{R}$ such that $|\varphi(t)| \ge \delta$ when $|t - c| \le \delta$. For fixed numbers a and b the set $\{\, \mathbf{x} \in \mathbb{R}^d \mid |\frac{\mathbf{u}\cdot\mathbf{x}-b}{a} - c| \le \delta \,\}$ has infinite measure when $d > 1$ and it follows that the function $\frac{1}{\sqrt{a}}\varphi(\frac{\mathbf{u}\cdot\mathbf{x}-b}{a})$ does not belong to $L^2(\mathbb{R}^d)$.

---

3. Suppose we are given m functions φ_j, $j = 1, \dots, m$ and n points $(\mathbf{x}_i, y_i)$ with $m > n$. How can one, using Lagrange multipliers find numbers c_j, $j = 1, 2, \dots, m$ such that $\sum_{j=1}^m c_j\varphi_j(\mathbf{x}_i) = y_i$ for all $i = 1, \dots, n$ and $\sum_{j=1}^m c_j^2$ is as small as possible?
Solution: Using Lagrange multipliers we see that we have to find the critical points of the function

$$F(c_1, \dots, c_m, \lambda_1, \dots, \lambda_n) = \frac{1}{2}\sum_{j=1}^{m} c_j^2 - \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{m} c_j\varphi_j(x_i) - y_i \right).$$

Differentiating with respect to c_j we get the equation

$$c_j - \sum_{i=1}^{n} \lambda_i\varphi_j(x_i), \quad j = 1, \dots, m.$$

Let $A(j,k) = \varphi_j(x_k)$ and denote by Y, C and Λ the column vectors with elements y_i, c_j and λ_i, respectively. Thus we can write the equation above in the form

$$C = A\Lambda.$$

On the other hand, the equations $\sum_{j=1}^m c_j\varphi_j(x_i) = y_i$ can be writen in the form

$$A^{\mathrm{T}}C = Y,$$

and hence we get

$$A^{\mathrm{T}}A\Lambda = Y \quad \text{and so} \quad \Lambda = (A^{\mathrm{T}}A)^{-1}Y,$$

so that

$$C = A(A^{\mathrm{T}}A)^{-1}Y.$$

C1. Write a matlab function `fferr` such that `[f,fp]=fferr(w,aux)` calculates the error and the derivative with respect to the weights and thresholds of a feed-forward neural network with dimensions, inputs, and outputs given in the vector `aux`.