HUT , Institute of mathematics — Gripenberg

Mat-1.196 Mathematics of neural networks

Exercise 1

15–22.1.2002

**1.** Why is it reasonable to call the function $\sigma(\mathbf{w}\cdot\mathbf{x}-\tau)$ a *ridge* function? Is this the case in dimension 1 as well?

*Solution:* If $\mathbf{x} = s\mathbf{v}+\mathbf{u}$ where $\mathbf{v}\perp\mathbf{w}$, then $\sigma(\mathbf{w}\cdot\mathbf{x}-\tau) = \sigma(\mathbf{w}\cdot\mathbf{u}-\tau)$, that is, this function is a constant on lines perpendicular to $\mathbf{w}$, and drawing the graph of the function $\sigma(\mathbf{w}\cdot\mathbf{x}-\tau)$ in the case where the dimension is 2 one gets a picture looking like a ridge, provided, of course, $\sigma$ is increasing on $(-\infty,a)$ and decreasing on $(a,\infty)$ for some $a$. In dimension 1 there are no directions perpendicular to $\mathbf{w}$ (except 0), so one does not really get a ridge in this case.

---

**2.** Show that the polynomials (of one variable) of degree at most $k$ are not dense in the spaces $\mathcal{C}([a,b])$ (continuous functions on $[a,b]$).

*Solution:* Suppose that this is not the case, so that the polynomials of degree at most $k$ are dense in $\mathcal{C}([a,b])$. Let $x_j = a + j\frac{b-a}{k+1}$, $j = 0,1,\ldots,k+1$. Let $f(x_j) = (-1)^j$ for $j = 0,1,\ldots,k+1$ and define $f(x)$ by piecewise linear interpolation at all other points of $[a,b]$. Then $f\in\mathcal{C}([a,b])$ and if the polynomials of degree at most $k$ are dense in $\mathcal{C}([a,b])$ then there is a polynomial $P_k$ so that $\sup_{x\in[a,b]}|P_k(x)-f(x)| < \frac{1}{2}$. But then $(-1)^j P_k(x_j) > 0$ for $j = 0,1,\ldots,x_{k+1}$ and it follows from the intermediate value theorem that $P_k$ has at least $k+1$ different zeros. But this is impossible and from this contradiction the claim follows.

---

**3.** Show that if $d\geq 1$ and $\sigma$ is a polynomial, then $S_d(\sigma)$ is not dense in $\mathcal{C}(\mathbb{R}^d)$ when $S_d(\sigma) = \operatorname{span}\{\,\mathbf{x}\in\mathbb{R}^d \mapsto \sigma(\mathbf{w}\cdot\mathbf{x}-\tau) \mid \mathbf{w}\in\mathbb{R}^d,\quad \tau\in\mathbb{R}\,\}$.

*Solution:* If $\sigma$ is a polynomial of degree at most $k$, then the function $x\mapsto\sigma(\mathbf{w}\cdot(x,0,\ldots,0)-\tau)$ is also a polynomial of degree at most $k$. Thus we see that if $f\in S_d(\sigma)$, then the function $g(x) = f((x,0,\ldots,0))$ is a polynomial of degree at most $k$ and since these polynomials are not dense in $\mathcal{C}(\mathbb{R})$ it follows that $S_d(\sigma)$ cannot be dense in $\mathcal{C}(\mathbb{R}^d)$.

---

**4.** Show that

$$d(f,g) = \sum_{j=1}^{\infty} 2^{-j}\frac{\|f-g\|_{\mathcal{B}^\infty(B_j(0)}}{1+\|f-g\|_{\mathcal{B}^\infty(B_j(0))}},$$

is a metric in the space $\mathcal{C}(\mathbb{R}^d)$ where $B_j(0) = \{\,\mathbf{x}\in\mathbb{R}^d \mid |x| < j\,\}$ and $\|f\|_{\mathcal{B}^\infty(K)} = \sup_{\mathbf{x}\in K}|f(\mathbf{x})|$.

*Solution:* It is clear that $0\leq d(f,g) = d(g,f) < \infty$ for all $f$ and $g\in\mathcal{C}(\mathbb{R}^d)$ and that $d(f,g) = 0$ if and only if $f = g$. Thus is remains to prove the triangle inequlity. If $f$, $g$, and $h\in\mathcal{C}(\mathbb{R}^d)$ and $j\geq 1$, then we clearly have

$$\|f-g\|_{\mathcal{B}^\infty(B_j(0)}\leq\|f-h\|_{\mathcal{B}^\infty(B_j(0)}+\|h-g\|_{\mathcal{B}^\infty(B_j(0)}.$$

Thus it suffices to show that if $a \le b + c$ where $a$, $b$ and $c$ are nonnegative numbers, then

$$\frac{a}{1+a} \le \frac{b}{1+b} + \frac{c}{1+c}.$$

Since the function $t \mapsto \frac{t}{1+t}$ is increasing, we have have

$$\frac{a}{1+a} \le \frac{b+c}{1+b+c},$$

and since a straightforward calculation shows that

$$\frac{b}{1+b} + \frac{c}{1+c} - \frac{b+c}{1+b+c} = \frac{(2+b+c)bc}{(1+a)(1+b)(1+b+c)} \ge 0,$$

the claim follows.