Solutions to Homework Set Three

ECE 271A
Electrical and Computer Engineering
University of California San Diego
Nuno Vasconcelos

1.

a) To minimize $f(\theta) = ||\mathbf{z} - \mathbf{\Phi}\theta||^2$ we compute the gradient

$$\nabla_\theta f = -2\mathbf{\Phi}^T(\mathbf{z} - \mathbf{\Phi}\theta)$$

and set it to zero, which leads to the standard least squares solution

$$\theta^* = \left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^T\mathbf{z}.$$

The matrix $\left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^T$ is referred to as the *pseudo-inverse* of $\mathbf{\Phi}$. To check that we have a minimum we compute the Hessian

$$\nabla^2_\theta f = 2\left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^T = 2\mathbf{\Phi}^T\mathbf{\Phi}$$

and use the theorem, which we already mentioned in the last homework set, that $\mathbf{A}$ is positive definite if and only if there is a non-singular matrix $\mathbf{R}$ such that $\mathbf{A} = \mathbf{R}^T\mathbf{R}$. (See e.g. *Linear Algebra and its Applications* by G. Strang, Harcourt Brace Jovanovic, 1988.) Since this is exactly the case of our Hessian, we only have to prove that $\mathbf{\Phi}$ is non-singular. We cannot really do this without knowing what the sample points x_i are but, assuming that all the x_i are different, this will hold.

b) Since, when x is known, $f(x, \theta)$ is a deterministic function of θ, and $\epsilon \sim \mathcal{N}(0, \sigma^2)$ it follows that

$$P_{Z|X}(z|x;\theta) = \mathcal{G}(x, f(x,\theta), \sigma^2)$$

Given an iid sample $\mathcal{D} = \{\mathcal{D}_x, \mathcal{D}_y\}$, the ML estimate of θ is then

$$\begin{aligned}
\theta^* &= \arg\max_\theta \sum_i \log P_{Z|X}(z_i|x_i;\theta) \\
&= \arg\max_\theta \sum_i -\frac{(z_i - f(x_i,\theta))^2}{2\sigma^2} - \frac{n}{2}\log(2\pi\sigma^2) \\
&= \arg\min_\theta \sum_i (z_i - f(x_i,\theta))^2 \\
&= \arg\min_\theta \sum_i (z_i - \phi_i\theta))^2 \qquad (1)
\end{aligned}$$

where $\phi_i = (1, x_i, \ldots, x_i^K)$, i.e. the i^{th} row of $\mathbf{\Phi}$. It follows that the ML solution is the one that minimizes $||\mathbf{z} - \mathbf{\Phi}\theta||^2$ and therefore the same as the least squares solution.

c) The only differ

Page 1 / 9