Suggestions, PS 7

5.4. **(a)** Recall natural splines are piecewise cubic on $[a, b]$ with knots at $\{\xi_1, ..., \xi_K\}$ which are restricted as linear in $[a, \xi_1]$ and in $[\xi_K, b]$. Show linearity in the first interval is equivalent to $\beta_2 = \beta_3 = 0$. In $[\xi_K, b]$ how does $f(X)$ simplify? Use this fact to get the remaining conditions.

(b) For the basis functions in (5.4), you can prove they are a basis for natural splines by showing (i) they are in the space, (ii) they span the space of cubic splines with knots at the points ξ_k, and (iii) that they are linearly independent. Why are these sufficient? For (i) recall the conditions in (a) above. For linear independence, why is it enough to show that each function in list (5.4) is linearly independent of the previous ones?

To show that each N_k is linearly independent of the previous $N_1, ..., N_{k-1}$, note $N_1 = 1\,, N_2 = x$ while

$$N_3(x) = d_1(x) - d_{K-1}(x) = \frac{(x-\xi_1)_+^3 - (x-\xi_K)_+^3}{\xi_K - \xi_1} - \frac{(x-\xi_{K-1})_+^3 - (x-\xi_K)_+^3}{\xi_K - \xi_{K-1}}\,;$$

show N_3 has a discontinuity in its third derivative at ξ_1, because of the term $(x-\xi_1)_+^3$. Since N_1 and N_2 do not have such discontinuities, it follows that N_3 cannot be a linear combination of N_1, N_2. Similarly, N_4 has a third derivative discontinuity at ξ_2, and so cannot be a combination of the previous functions, none of which have derivative discontinuities at ξ_2. Continue this way, until noting that $N_{K-1} = d_{K-3} - d_{K-1}$ has a third derivative discontinuity at ξ_{K-3}, while none of the previous N_k do, and N_K has a third derivative discontinuity at ξ_{K-2}, while none of the previous ones do. Thus conclude each function is linearly independent of the previous ones. How does this prove linear independence?

Finally, you need to show these functions span your natural spline space (call it V). First it suffices to show V is indeed K dimensional, since from linear algebra we know if vector space V has K linearly independent vectors and a spanning set of K vectors (not necessarily the same), then it is K dimensional and every linearly independent set of K vectors is spanning.

Show V is K dimensional look at the (larger) space U of *all* cubic splines on the mesh $\{\xi_k\}_{k=1}^K$ that are linear only in the *first* interval. Analogously to general cubic splines, a simple basis is $\{\phi_0 = 1, \phi_{-1} = x\} \cup \{\phi_k = (x-\xi_k)^3\}_{k=1}^K$ (why?). Show this is a $K+2$ dimensional space, but it is strictly larger than V (our space of natural splines).

However, if $f \in V$ is to be written $f = \sum_{k=-1}^{K} a_k \phi_k(x)$, show we must just additionally require f be linear in the last interval, i.e. in $[\xi_K, 1]$ (and hence everywhere-why?), we have $a_0 + a_{-1}x + \sum_{k=1}^{K} a_k (x-\xi_k)^3 = A_0 + A_1 x + A_2 x^2 + A_3 x^3$, with $A_2 = A_3 = 0$.

Show this gives two necessary linear relationships between the $\{a_k\}_{k=-1}^{K}$ and they are independent, i.e., one does not determine the other.

Thus show that V is in one to one correspondence with the set V_1 of all sets of coefficients $\{a_{-1}, a_0, \ldots, a_K\}$ satisfying two independent linear relations. Show this forms a subspace of the larger $K+2$ dimensional space U_1 consisting of *all* sets of such coefficients. Show vectors in space U_1 (as vector arrays of numbers) satisfing two independent equations (defining V_1) must forn a K dimensional vector space (in the sense of linear algebra). Thus conclude the set V of natural cubic splines must be K dimensional.

5.7. (a) You can show by breaking up the left side that

$$\int_a^b g''(x)h''(x)dx \ =$$

$$= \sum_{j=1}^{K-1} - \int_{x_j}^{x_{j+1}} g'''(x)h^{'}(x)dx + \sum_{j=1}^{K-1} \{g''(x_{j+1})h^{'}(x_{j+1}) - g''(x_j)h^{'}(x_j)\}.$$

(why is the first term $\int_a^{x_1} g'''(x)h^{'}(x)dx = 0$?) Show after cancellations the second sum equals

$$g''(x_K)h^{'}(x_K) - g''(x_1)h^{'}(x_1) = 0$$

(what must $g''(x_1)$ be?). You can IBP again obtaining

$$= \ - \sum_{j=1}^{K-1} g'''(x)h(x)|_{x_j}^{x_{j+1}} \qquad (1)$$

(why is $g'''' = 0$)? You need to be careful defining the above sum $-$ remember that $g'''(x_j^+) \neq g'''(x_j^-)$, where + means right hand limit. Note $g'''(x_{j+1}^-) = g'''(x_j^+)$ (why?), to obtain from (1)

$$- \sum_{j=1}^{N-1} g'''(x_j^+)[h(x_{j+1}) - h(x_j)] = 0\ .$$

Why is $h(x_j) = 0$?

(b) You can try using

$$\int_a^b \tilde{g}''(x)^2\, dx = \int_a^b [\tilde{g}''(x) - g''(x) + g''(x)]^2\, dx.$$

(c) If $g(x)$ is differentiable, consider the Lagrangian penalty

$$\mathcal{L}(g) = \sum_{j=1}^{N}(y_i - g(x_i))^2 + \lambda\int_a^b g''(x)^2\, dx.$$

Assume $\mathcal{L}$ has a minimizing function $g_1 \in W^{(2)}$ (functions with two continuous derivatives), and define $z_i = g_1(x_i)$. Show now that the cubic spline solution $g(x)$ defined in (a) makes both the first *and* the second terms the smallest possible, among *all* functions $\widetilde{g}(x) \in W^{(2)}$ that interpolate the points (x_i, z_i) (i.e. that go through these same points).