5. *Semi-parametric Linear Model.* Consider an additive model \( y_i = x_i^T \beta + f(z_i) + \epsilon_i, \quad i = 1, \ldots, n. \) Here \( x_i \) is a vector of \( p \) predictors (including the element 1), and \( z_i \) is an additional confounding predictor. You plan to fit this model by penalized least squares, using a smoothing spline to control the roughness of \( f \) (\( \beta \) will be unpenalized).

(a) Write down the penalized least-squares criterion for fitting this model.

(b) Show that the minimizing \( \hat{\beta} \) and \( \hat{f} \) satisfy the following pair of backfitting equations:

\[
\begin{align*}
X \hat{\beta} &= H(y - \hat{f}) \\
\hat{f} &= S_\lambda(y - X \hat{\beta}),
\end{align*}
\]

where \( f \) is the vector of \( n \) fitted values for the function \( f \), and \( S_\lambda \) is the appropriate smoothing spline operator matrix. What is \( H \)?

(c) Show that you can solve these equations explicitly for \( \hat{\beta} \), and hence for \( \hat{f} \) as well.

(d) The smoothing spline operation \( S_\lambda r \) for any vector \( r \) can be computed in \( O(n) \) operations. Discuss how you might organize the computations in the previous item in an efficient manner. What is the order of computations for your entire solution.

6. (a) ESL 9.4. Start off by showing that this is the system of equations that corresponds to backfitting. What if each of the \( S_j \) corresponds to a ridged regression:

\[
S_j = B_j (B_j^T B_j + \lambda I)^{-1} B_j^T;
\]
Describe in compact form the optimization problem that is being solved, the solution, and that the system of equations given in the problem is equivalent to this solution.

7. ESL 9.7(a) *Analysis of the ozone data, available on the ESL website.* In Splus there is a function `gam`; in R there is a contributed package called `gam`. Explore the model using either the informal anova techniques discussed in class, or any techniques you choose. Decide if all the terms are needed, and how each should be fit.