# Linear Models sheet 2

1. For each of the following simple linear models, write down the appropriate $X$ matrix, and derive least squares estimators for the parameters, both from the general matrix formula $\widehat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \boldsymbol{Y}$ and from first principles, by differentiation. Check that your answers agree!

    (a) Linear regression through the origin: $Y_i = \beta x_i + \epsilon_i, i = 1, 2, \ldots, n$.

    (b) Linear model with one factor: $Y_{ij} = \alpha_i + \epsilon_{ij}, j = 1, 2, \ldots, n_i; i = 1, 2, \ldots, k$.

2. In the general linear model $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, show that the least squares estimator $\widehat{\boldsymbol{\beta}}$ satisfies the normal equations $(X^T X)\widehat{\boldsymbol{\beta}} = X^T \boldsymbol{Y}$, not by the algebraic method in the notes, but by partial differentiation with respect to $\beta_j, j = 1, 2, \ldots, p$.

3. Which of the $X$ matrices in section 1 of the notes is *not* of full rank? Which others might or might not be, depending on the values of the $x$ variables?

4. Prove the facts about $H = X(X^T X)^{-1} X^T$ stated on page 19 of the notes: $H^T = H$, $H^2 = H$ and $(I - H^T)(I - H) = (I - H)$. [Hint: you may assume that for any matrices for which *both sides of* these equalities make sense: $(AB)^T = B^T A^T$, $(AB)^{-1} = B^{-1} A^{-1}$.]

5. Prove that if $A$ and $B$ are suitably conformal matrices, then $(AB)^T = B^T A^T$, and $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ (where $\mathrm{tr}(\cdot)$ means the trace of a matrix, i.e. the sum of the diagonal entries).

6. Prove the facts quoted on page 24 of the notes:

    for any random $n$-vector $\boldsymbol{Y}$, and any constant $m \times n$ matrix $A$, $E(A\boldsymbol{Y}) = AE(\boldsymbol{Y})$ and $\mathrm{var}(A\boldsymbol{Y}) = A\mathrm{var}(\boldsymbol{Y})A^T$.

    and that on page 27:

    for any random vector $\boldsymbol{Y}$ with $E(\boldsymbol{Y}) = \boldsymbol{\mu}$ and $\mathrm{var}(\boldsymbol{Y}) = V$, and any constant matrix $A$, $E(\boldsymbol{Y}^T A\boldsymbol{Y}) = \boldsymbol{\mu}^T A\boldsymbol{\mu} + \mathrm{tr}(AV)$.

    by expressing both sides in terms of the individual components of the matrices and vectors, and applying basic facts about the mean and variance of a sum of random variables.

7. Find an expression for the diagonal entries $h_{ii}$ of the hat matrix $H$ for simple linear regression. (You can use either the $Y_i = \alpha + \beta x_i + \epsilon_i$ or $Y_i = \alpha^\star + \beta(x_i - \overline{x}) + \epsilon_i$ versions of the model – why does it make no difference which?) Recall that $\mathrm{var}(e_i) = \sigma^2(1 - h_{ii})$: for which values of $x_i$ is this particularly large or small? Does this make sense intuitively?

8. Analyse the `rubber` data set (available in `LM.RData` either from the unit web page, or in the undergrad computing lab). Try several different linear models for the response variable `Abrasion` in terms of `Hardness` and `Tensile`. Examine the residual sums of squares and degrees of freedom of each model that you try, and study the graphics produced by the `plot` command, applied to the output of the run of `lm`.

    (Brief information about data sets in the `LM.RData` workspace can be obtained by typing, e.g., `desc(rubber)`).