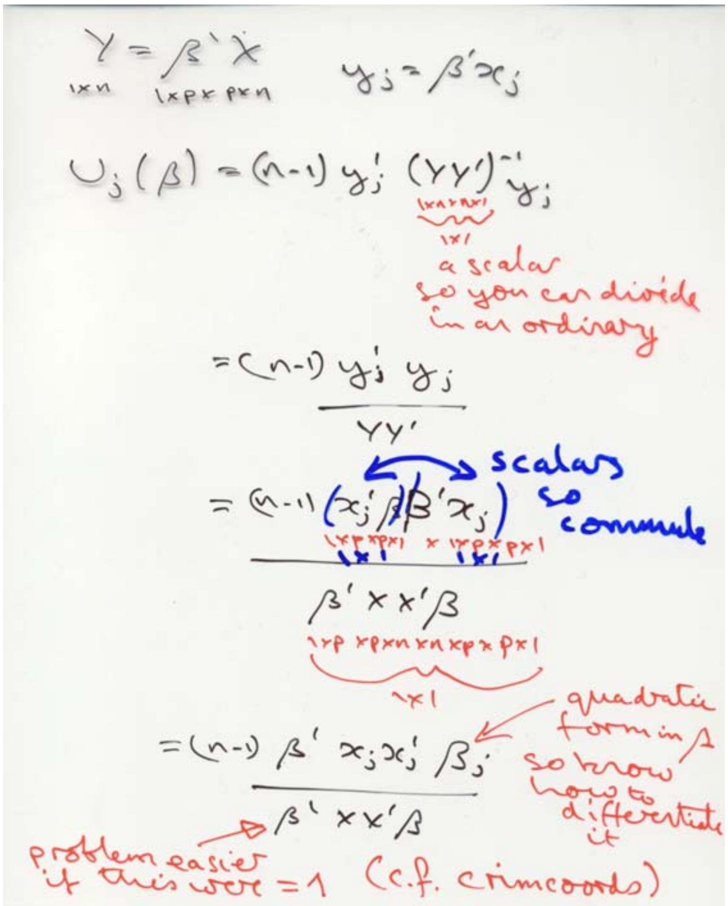
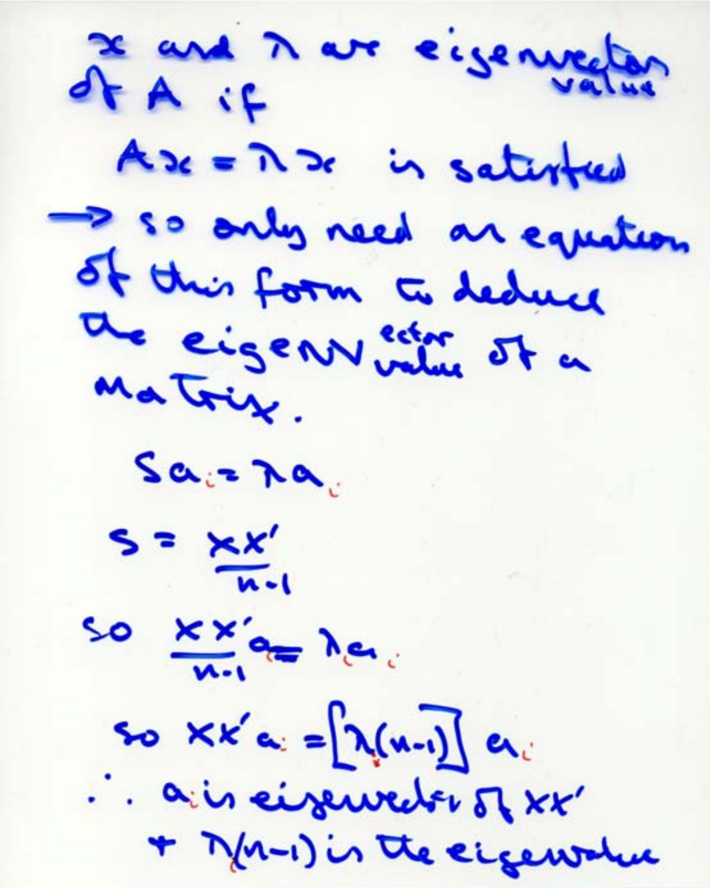
Notes from lecture 29/10/07 & 01/11/07

These are copies of the OHP transparencies from the lectures on Monday 29/10/07 & Thursday 01/11/07 relating to Exercises 1, Q1 and Task Sheet 4, Q1.



The first step in manipulating long expressions involving matrices and vectors is to annotate each element with its dimensions. The dimensions (e.g. nxp) have to match up rather like playing dominoes (sorry if this is a UK cultural feature), i.e. the adjacent elements to one which is nxp must be something such as sxn and pxt so you have sxnxnxpxpxt sp the whole string is sxt. Then it is useful to identify a string which is of the form 1x.........x1 where there may be several elements in between. Usually the first element of such a string will be the transpose of a vector and the last will also *be a vector. The importance of this is that the complete string (e.g. $\boldsymbol{\beta}'x_jx_j'\boldsymbol{\beta}$) is a scalar and so can move around, i.e. it is an ordinary number so you can move it to the front or if it comes inside an inverse like $(\boldsymbol{\beta}'x_jx_j'\boldsymbol{\beta})^{-1}$ then you can turn the expression into a fraction and divide. This was the first step above. Then the same task was applied to the top line of the fraction (after substituting in the expression for $y_j=\boldsymbol{\beta}'x_j$) and noting that these also are scalars so we can shuffle the order of multiplication (not allowed with general matrices but these are ordinary numbers). Then this looks very like the expression for F_1 on P143, with $B = x_jx_j'$, $W=XX'$ and $a=\boldsymbol{\beta}$. So, you can just repeat the mathematics on P145 with these modifications. This should take you through (i) to (iii) and then (iv) is covered in the notes on eigen analyses of special matrices posted last week.*

This question is undeniably hard for people coming to matrices for the first time and when you look at past examination questions I [personally] do not set questions such as this where all I will discover is that half the people have trouble with the mathematics and I don't discover how much they know about statistics, so exams are more like Q2. The reason I have nevertheless included it is because it has an application which is of considerable importance in detection and assessment of multivariate outliers which is described in the article attached to the message in the Scatter plots & Outliers thread on 23rd October at 8:25 and which is on the PAS470 page.



x and λ are eigenvector/value of A if

$Ax = \lambda x$ is satisfied

→ so only need an equation of this form to deduce the eigenvector/value of a matrix.

$Sa_i = \lambda a_i$

$S = \frac{XX'}{n-1}$

so $\frac{XX'}{n-1} a_i = \lambda a_i$

so $XX' a_i = [\lambda(n-1)] a_i$

∴ a_i is eigenvector of XX' + $\lambda(n-1)$ is the eigenvalue

Re Q1 of task sheet 4, the important feature is that if a vector x and a scalar λ satisfy the eigen equation then they are the eigenvector and value for that matrix. Nothing more needs to be proved (just as if x satisfies an equation of a polynomial = 0 then it is a root of that polynomial).