Multivariate Methods

Homework #1 (2018-2019), Questions

Q1: Another Lagrange Multiplier application

(As you may well know) the entropy of a discrete distribution is given by $H(\vec{p}) = -\frac{1}{\log 2}\sum_i p_i \log p_i$.
Consider a discrete distribution in which p_n is the probability of drawing a token of "value" n, where $n = 0,1,2,3,\cdots$. Find the distribution $\vec{p}$ that maximizes $H(\vec{p})$ subject to the constraint that the average value is equal to A, i.e., that $\sum_{i=0}^{\infty} ip_i = A$.

Q2: Regression and "cross-correlation analysis" (from MVAR1415)

Consider the standard regression scenario described in the class notes, pages 1-2. That is, there are n observations, $y_1,\ldots,y_n$, and p regressors, where the typical regressor $\vec{x}_j$ is a column $x_{1,j},\ldots x_{n,j}$, and the set of p regressors forms a $n\times p$ matrix X, and we seek a set of p coefficients $b_1,\ldots,b_p$, the $p\times 1$ matrix B, for which $\left|Y - XB\right|^2$ is minimized.
Now let's assume that the regressors $\vec{x}_j$ are orthonormal. For example, we're doing spatial receptive field analysis. Here $x_{i,j}$ corresponds to the luminance presented on the ith trial in pixel j, and we've designed our stimuli so that, over the entire stimulus sequence, $\sum_{i=1}^{N} x_{i,j}x_{i,k} = 0$ if $j \neq k$, and $\sum_{i=1}^{N} x_{i,j}x_{i,j} = 1$.
How does this simplify the computation of the regressors B?