**Question A.**  
You have two separate, intercept-free OLS regressions:

* Y1∼X1 on M observations yields coefficient β1.
* Y2∼X2 on N observations yields coefficient β2.

You then pool the data (Y=[Y1;Y2], X=[X1;X2], no intercept) and fit Y∼X to obtain β.

1. **No intercept case**  
   Given β1 and β2, what is the set of all possible values of β?
2. **With intercepts**  
   Repeat the above, but now each regression (on (X1,Y1), (X2,Y2), and on the pooled data) includes an intercept. Denote the slopes by β1,β2,β. Assuming the feasible β form a continuous interval, what is that interval in terms of β1 and β2?
3. **Gaussian data approximation**  
   Returning to the intercept-free setup, assume all (X,Y) pairs are drawn i.i.d. from a zero-mean bivariate normal. Given β1 and β2, what is your best estimate of the pooled slope β? You may invoke reasonable approximations.

**Question B.**  
Let X1 be an N×k matrix of regressors and X2 an N×(f−k) matrix of additional candidates. You have already fit Y∼X1 (via QR decomposition), producing residuals e.

1. **Best single addition**  
   Describe an efficient algorithm for selecting which single column of X2 will, if added to X1, yield the largest drop in residual sum of squares.
2. **Updating the fit**  
   Once you’ve chosen that best column from X2, derive an efficient procedure to update the regression coefficients for the model Y∼[X1,chosen column], ideally reusing the existing QR factorization of X1.