Quiz 8. Shrinkage Methods (cont.), PCA and PLS

Name: Lisa Liubovich                    Attempt (circle one):    BEFORE    (AFTER)

*In Q.1 and Q.4, circle **all** correct answers. <u>There may be more than one correct choice</u>.*

1.  Principle Component Analysis (PCA) methods are used for:

    (a) Data augmentation     (b) Dimension reduction     (c) Reducing sample size

2.  In linear regression, one may want to use the principal components regression or the partial least squares (PLS) in order to:

    (a)  achieve better prediction accuracy (such as MSE) by reducing the model flexibility.

    (b)  reduce multicollinearity

    (c)  reduce variance of estimated regression parameters.

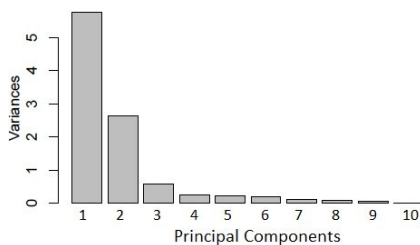3.  Let $Z_1$ and $Z_2$ be the 1st and the 2nd principal components vectors. $Z_1$ is _____ to $Z_2$.

    (a) parallel                    (b) orthogonal.

4.  Let Z1 and Z2 be the 1st and the 2nd principal components vectors. $Var(Z_1)$ _____ $Var(Z_2)$.

    (a)       >                               (b)       <

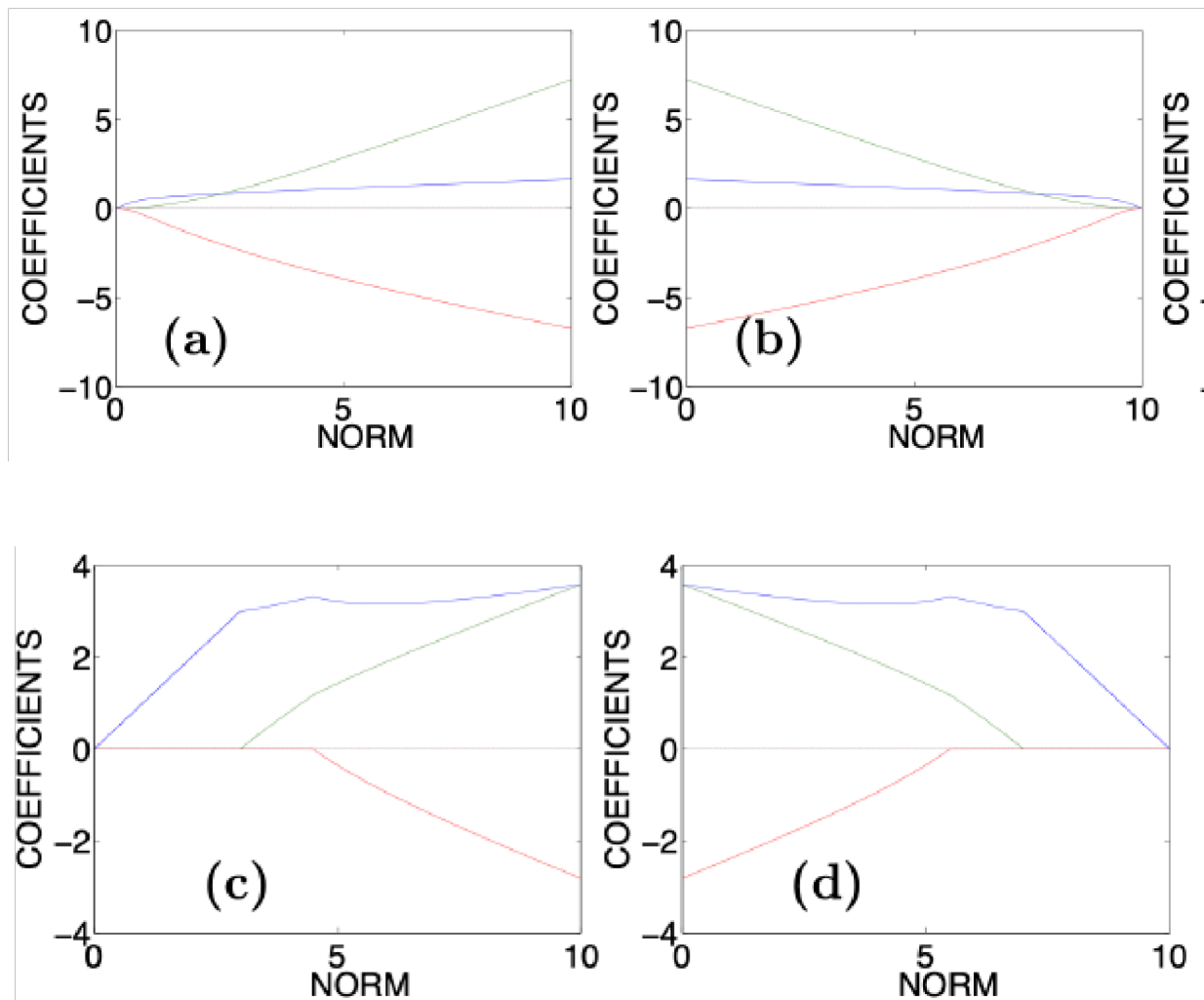5.  **Stat-427.** What is the purpose of each plot?
    **Stat 627.** Pick ONE of the following plots from PCA and Principle Component Regression. What is the purpose of the plot?



This plot is a scree plot, which helps us decide how many principal components to use in an ad hoc way by looking at where the variance drops off for each additional principal component (known as the elbow).

Stat 427/627 Statistical Machine Learning        Quiz 8                    Jun Lu. American University
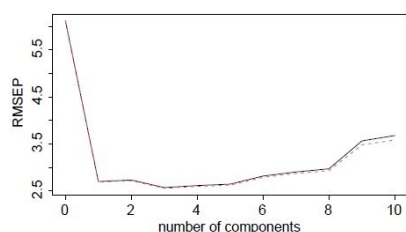
6. **Stat-627 Only**. Among the following plots:



(a) Which plot can represent the ridge regression coefficient?
   Plots A because it shows a distinct smooth shrinkage towards zero without actually reaching zero

(b) Which plot can represent the lasso coefficient?
   Plot D because it shows a sharp change of the coefficients dropping to 0