

1. Let X be a design matrix for a linear regression model. The problem of multicollinearity is arises when:	4 / 4 points	5. Which of the following metrics can be used to diagnose multicollinearity?	4 / 4 points
<div>The regression of the response on a subset of other predictors produces a high R^2 value.</div> <div>One column of the design matrix X is a near linear combination of other columns.</div> <div>The regression of one predictor on several others produces a high R^2 value.</div> <div>The matrix $(X^T X)$ is close to singular (singular means not invertible).</div> <div>One column of the design matrix X is a near constant multiple of another.</div>		<div>Variance inflation factors.</div> <div>Pairwise correlations between response variables.</div> <div>The condition number of the matrix $(X^T X)$.</div> <div>Pairwise correlations between predictors.</div> <div>Pairwise correlations between residuals.</div>	
2. A matrix X is ill-conditioned when $X^T X$ cannot be computed with reliable accuracy.	3 / 3 points	6. Let $\hat{\beta}_j$ be a least squares estimator of the parameter β_j . The variance of $\hat{\beta}_j$ is impacted by:	4 / 4 points
<div>True</div> <div>False</div>		<div>The correlation between the j^{th} predictor and all other predictors.</div> <div>The correlation between the response and the j^{th} predictor.</div> <div>The variability of the j^{th} predictor measurements.</div> <div>The variance of the error terms in the regression model.</div>	
3. Suppose that data are collected to predict the daily risk of a wildland fire in a particular region based on a set of predictors. The predictors include daily high air temperature in the region, daily number of lightning strikes in the region, the daily number of times thunder was heard in the region, and the daily number of registered campers in the region.	3 / 3 points	7. Variance inflation factors, the condition number for $X^T X$, and pairwise predictor correlations will always agree in their diagnosis of multicollinearity.	3 / 3 points
<div>The daily number of lightning strikes in the region and the daily number of times thunder was heard in the region are highly correlated. This high correlation is likely due to:</div> <div>Statistical misspecification</div> <div>Lack of information</div>		<div>True</div> <div>False</div>	
4. Suppose that data have been collected to predict the miles per gallon for a given vehicle based on a set of predictors. The predictors include the weight of the vehicle, the number of engine cylinders, the number of carburetors, and whether the car was automatic or manual transmission.	3 / 3 points	8. Methods to rid a model of multicollinearity include:	4 / 4 points
<div>In the data, the vast majority of automatic transmission vehicles are light ($< 2,500$ lbs), and the vast majority manual transmission vehicles are heavy ($> 2,500$ lbs). Suppose that, in the entire population of cars being studied, however, there are many automatic transmission vehicles above 2,500 lbs and many manual transmission vehicles under $< 2,500$ lbs.</div> <div>The high correlation between transmission type and weight is likely due to:</div> <div>Statistical misspecification</div> <div>Lack of information</div>		<div>Alternative estimation methods.</div> <div>Principal component analysis</div> <div>Changing the response variable.</div> <div>Removing predictors.</div> <div>Pairwise correlations between residuals.</div>	