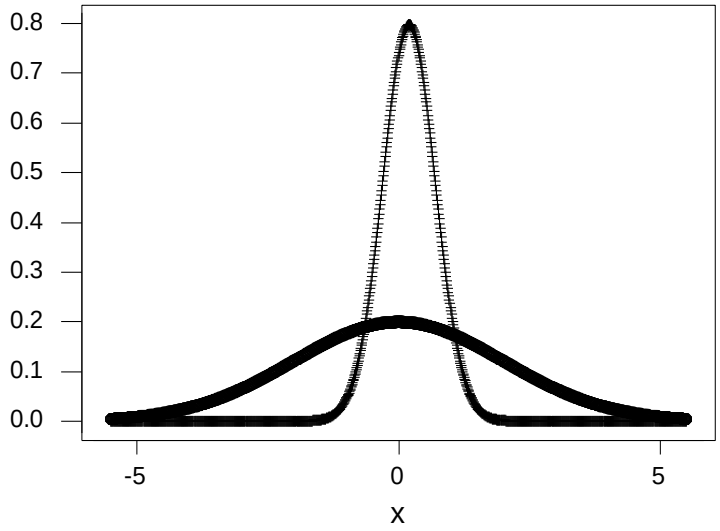
Remedial Measures for Multicollinearity

1. Making predictions is not a problem as long as we use the predictor variables that follow the same pattern of multicollinearity
2. Use of centered data for predictor variables when using polynomial regression and other regressions with interactions
3. Dropping one or more predictors
4. Add some cases that may break the pattern of multicollinearity
5. Use different data sets to estimate different coefficients
6. Principal component type analysis

Ridge Regression:

Idea of biased regression:



Bias vs. Variance. Which one is more critical?

Mean Squared Error =
$E\{b^R - \beta\}^2 = Var(b^R) + (E[b^R]-\beta)^2$

When b^R is unbiased then the second term drops out.

In estimating $\beta=0$, we have
In this plot, $E(b^B) = .2$, $Var(b^B) = .25$ and
Mean Squared Error=.29
$E(b^U)=0$, $Var(b^U)=4.0$ MSE =4.0
Mean Squared Error=4.0

Ridge Estimators

Using Correlation Transforms we have,

$r_{XX}\mathbf{b}=r_{Yx}$

Idea is to use a small biasing constant c and find bR as a solution to,
$(r_{XX}+c\mathbf{I})\,\mathbf{b}^{\mathbf{R}}=r_{Yx}$

So that $\mathbf{b}^{\mathbf{R}}$ are our standardized ridge coefficients.

Idea of Ridge Regression:

To make the correlation matrix r_{xx} more stable add a biasing constant to the diagonal and invert the biased but more stable matrix.

Choice of Biasing Constant, c:

Results:

1. As c increases, bias increases, variance of the betas decreases
2. There always exists a c for which the total MSE for ridge regression is SMALLER than that for OLS.
3. There are no hard and fast ways of finding c.

Finding c:
Often found using the ridge trace and variance inflation factors

- Ridge Trace:

Simultaneous plot of the values of the (p-1) estimated ridge standardized regression coefficients for different values of c between o and 1.

- These may fluctuate widely form each other if c is slightly changed from 0. Gradually this steadies down and the regression coefficients tend to move towards 0.
- VIF also falls as c changes
- Want c to be the value when regression coefficients have steadied itself and VIF is small.

Comments on ridge regression:

1. Normal equations

$(1+c)b_1^R + r_{12}b_2^R + \ldots + r_{1,p-1}b_{p-1}^R = r_{y1}$

$\ldots$

$r_{p-1,1}\, b_1^R + r_{p-1,2}b_2^R + \ldots + (1+c)\, b_{p-1}^R = r_{Yp-1}$

2. VIF can be defined the same way as in OLS
3. R^2 for ridge regression can be defined the same way as in OLS
4. The ridge regression estimates tend to be stable, in the sense that they are little affected by small changes in data.
5. All inference is only approximate