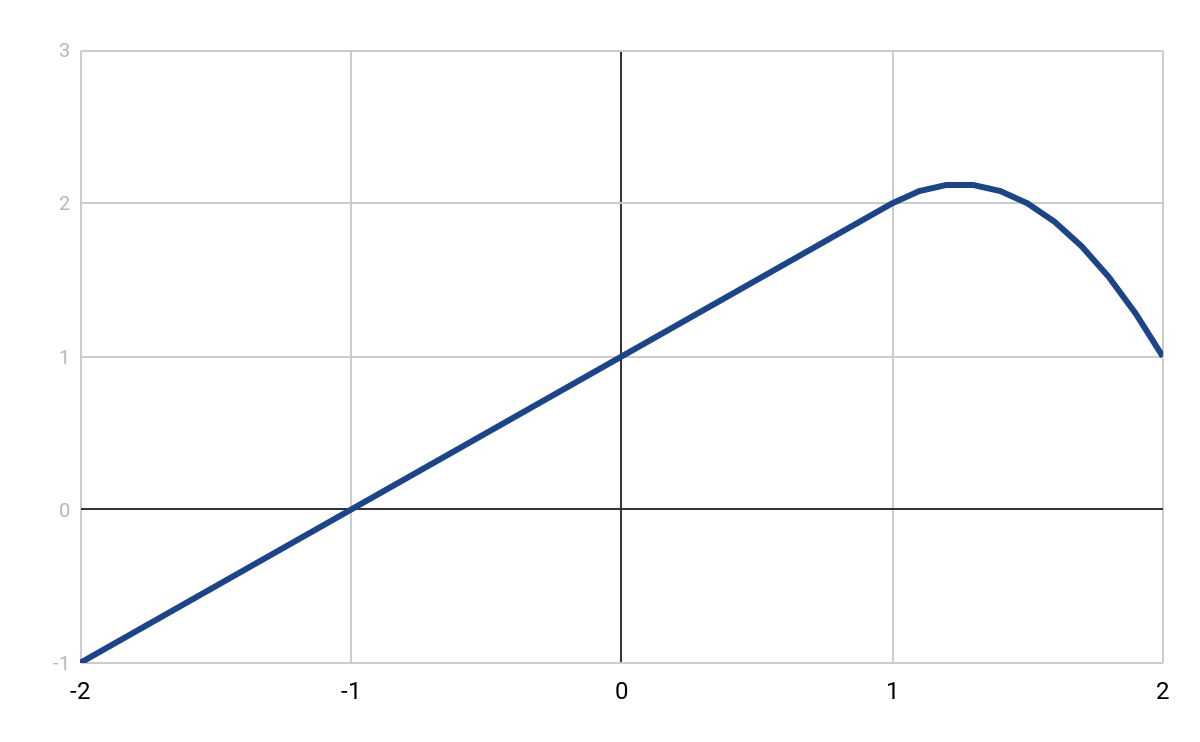
3. Suppose we fit a curve with basis functions b1(X) = X, b2(x) = (X-1)^2I(X>=1). We fit the linear regression model and obtain coefficient estimates of 1, 1, and -2 for B0, B1, and B2 resp. Sketch the estimated curve between X=-2 and 2.

**Answer**: Function takes the form of y = 1 + X -2(X-1)^2I(X>=1)



5. Consider two curves, g1 and g2, defined as

,

1. As , will g1 or g2 have the smaller training RSS?

**Answer**: g2 has a larger constraint (4th derivative), which means it is more flexible than g1. Therefore, g2 will fit the training data better and have a lower RSS.

1. Above, but test RSS?

**Answer**: Because g1 is less flexible, it won’t overfit the data as much, and will likely have a lower test RSS.

1. For , will g1 or g2 have the smaller training and test RSS?

**Answer**: When lambda approaches 0, the penalty term also approaches 0, which means the left part of the equation is all that remains. Since g1 and g2 have the same left hand portion, the RSS will be the same for both.