

The Bias-Variance Tradeoff for KNN

Use the following as your training data set:

And the following example of how to plot a function of k (with the training data fixed).

$$f(x) = -9.3 + 2.6x - 0.3x^2 + .01x^3$$

We can visualize the training data, the fitted model in blue (say, with $k = 3$) and the true model in gold.

```
knn <- function(x, k, training_data) { # from week 3 slides
  n <- length(x)
  y_hat <- rep(NA, n)
  for (i in 1:n) {
    dists <- abs(x[i] - training_data$x)
    neighbors <- order(dists)[1:k]
    y_hat[i] <- mean(training_data$y[neighbors])
  }
  y_hat
}

x_seq <- seq(.5, 12.5, length.out = 300)
y_hat <- knn(x = x_seq,
            k = 3,
            training_data = df_train)

f <- function(x) {
  f(x) = -9.3 + 2.6 * x - 0.3 * x^2 + .01 * x^3
}

df_lines <- tibble(x = rep(x_seq, 2),
                  y = c(f(x_seq), y_hat),
                  type = rep(c("f", "f_hat"), each = length(x_seq)))

ggplot(df_train, aes(x = x, y = y)) +
  geom_point() +
  theme_bw() +
  geom_line(data = df_lines, aes(col = type))
```

$$E \left[(y - \hat{f}(x))^2 \right] = \text{Var}(\hat{f}(x)) + \left[E(f(x) - \hat{f}(x)) \right]^2 + \text{Var}(\epsilon)$$

Variance

The variance term captures the degree to which the model, $\hat{f}(x)$ will vary from one data set to the next. This term can be better understood by replacing $\hat{f}(x)$ with the expression for the KNN model and use the properties of variance to simplify.

$$\begin{aligned}
\text{Var}(\hat{f}(x)) &= \text{Var}\left(\frac{1}{k} \sum_{x_i \in \mathcal{N}(x)} y_i\right) \\
&= \frac{1}{k^2} \sum_{x_i \in \mathcal{N}(x)} \text{Var}(y_i) \\
&= \frac{1}{k^2} \sum_{x_i \in \mathcal{N}(x)} \sigma^2 \\
&= \frac{1}{k^2} k \sigma^2 \\
&= \frac{\sigma^2}{k}
\end{aligned}$$

Line 3 is true because $\text{Var}(y_i) = \text{Var}(f + \epsilon) = \text{Var}(f) + \text{Var}(\epsilon) = 0 + \sigma^2$. This final expression looks familiar: it's the variance of a sample mean. In the KNN model, that's exactly what $\hat{f}(x)$ is: a sample mean of k y_i values.

We can write the variance as a function of k (and σ^2) as follows.

```
var_f_hat <- function(k, sigma) {
  sigma^2 / k
}
```

Bias

The bias is the amount by which the expected fitted model diverges from the true model. To better understand this term, we start by substituting in the form of the KNN model and apply the properties of expected value.

$$\begin{aligned}
E(f(x) - \hat{f}(x)) &= E(f(x)) - E(\hat{f}(x)) \\
&= f(x) - E\left(\frac{1}{k} \sum_{x_i \in \mathcal{N}(x)} y_i\right) \\
&= f(x) - \frac{1}{k} \sum_{x_i \in \mathcal{N}(x)} E(y_i) \\
&= f(x) - \frac{1}{k} \sum_{x_i \in \mathcal{N}(x)} f(x_i)
\end{aligned}$$

This says that the bias at any point x is the value of f at that point minus the average of the k values of the true regression function evaluated at the points x_i in the training data.

The bias term is more difficult to discuss generally because it deals with the true regression function f . This is not known in practice, so the sort of analysis that we are conducting presently has to be couched as, “If the true function was *this*, then the bias-variance trade-off would look like *that*”.