If our objective is to keep the mean squared estimation error small, then the best possible estimator is the conditional expectation.

But sometimes the conditional expectation is hard to calculate.

Maybe we're missing the details of the various probability distributions.

Or maybe we have the distributions that we need but the formulas are complicated.

After all, the conditional expectation can be a complicated non-linear function of the observations.

For this reason, we may want to consider an estimator that has a simpler structure, an estimator that is a linear function of the data.

And then, within this class of estimators, find the one that results in the smallest possible mean squared error.

In this lecture we will formulate this linear least squares estimation problem and then solve it.

We will see that the solution is given by a simple formula that involves only the means, variances, and covariances of the random variables involved.

Because of the simplicity of the method, linear estimators are used quite often, especially in systems where estimates need to be computed quickly in real time as observations are obtained.

We will look into some of the mathematical properties of the linear least mean squares estimator and the associated mean squared error, revisit an example from the previous lecture, and finally close with some comments on the ways that this estimator can be used.