

We have seen so far two ways of estimating an unknown parameter. We can use the maximum a posteriori probability estimate, or we can use the conditional expectation. That is, the mean of the posterior distribution. These were, in some sense, arbitrary choices.

How about imposing a performance criterion, and then finding an estimate which is optimal with respect to that criterion? This is what we will be doing in this lecture. We introduce a specific performance criterion, the expected value of the squared estimation error, and we look for an estimator that is optimal under this criterion.

It turns out that the optimal estimator is the conditional expectation. And this is why we have been calling it the least mean squares estimator. It plays a central role because it is a canonical way of estimating unknown random variables. We will study some of its theoretical properties, and we will also illustrate its use and the associated performance analysis in the context of an example.