Bias and **variance** measure two different sources of error in an estimator.

Bias measures the expected deviation from the true value of the function or parameter.

Variance on the other hand, provides a measure of the deviation from the expected estimator value that any particular sampling of the data is likely to cause.

Common Trade-off approaches

1. Cross-validation
2. Compare the Mean Squared Error (MSE) of the estimates

Proof of variance and bias relationship [edit]

$$
\begin{aligned}
\mathrm{MSE}(\hat{\theta}) &= \mathrm{E}_\theta\left[(\hat{\theta}-\theta)^2\right] \\
&= \mathrm{E}_\theta\left[\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]+\mathrm{E}_\theta[\hat{\theta}]-\theta\right)^2\right] \\
&= \mathrm{E}_\theta\left[\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)^2+2\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)+\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)^2\right] \\
&= \mathrm{E}_\theta\left[\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)^2\right]+\mathrm{E}_\theta\left[2\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)\right]+\mathrm{E}_\theta\left[\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)^2\right] \\
&= \mathrm{E}_\theta\left[\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)^2\right]+2\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)\mathrm{E}_\theta\left[\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right]+\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)^2 && \mathrm{E}_\theta[\hat{\theta}]-\theta=\text{const.} \\
&= \mathrm{E}_\theta\left[\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)^2\right]+2\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)\left(\mathrm{E}_\theta[\hat{\theta}]-\mathrm{E}_\theta[\hat{\theta}]\right)+\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)^2 && \mathrm{E}_\theta[\hat{\theta}]=\text{const.} \\
&= \mathrm{E}_\theta\left[\left(\hat{\theta}-\mathrm{E}_\theta[\hat{\theta}]\right)^2\right]+\left(\mathrm{E}_\theta[\hat{\theta}]-\theta\right)^2 \\
&= \mathrm{Var}_\theta(\hat{\theta})+\mathrm{Bias}_\theta(\hat{\theta},\theta)^2
\end{aligned}
$$

But in real modeling case, MSE could be described as the addition of model variance, model bias, and irreducible uncertainty. According to the relationship, the MSE of the estimators could be simply used for the efficiency comparison, which includes the information of estimator variance and bias. This is called MSE criterion.