

Descomposición del Error: Sesgo-Varianza

Derivación Matemática

July 24, 2025

Contexto del Problema

El objetivo es entender la naturaleza del error de un modelo en el aprendizaje supervisado. Partimos de los siguientes supuestos:

- El verdadero modelo que genera los datos es $y = g(x) + \epsilon$, donde el ruido ϵ es una variable aleatoria con media cero ($\mathbb{E}[\epsilon] = 0$) y desviación estándar σ . Por lo tanto, la varianza del ruido es $\text{Var}(\epsilon) = \sigma^2$.
- Aprendemos un predictor $\hat{y} = f(x; D)$ al que entrenamos en un conjunto de datos específico $D = \{(x_i, y_i)\}$. La forma del predictor f depende del conjunto de entrenamiento D .

Buscamos descomponer el error de generalización esperado para un nuevo punto de datos x .

Derivación del Error Cuadrático Esperado

El error de generalización se define como el valor esperado del error cuadrático, donde la esperanza se toma sobre todos los posibles conjuntos de entrenamiento D y sobre el ruido ϵ .

$$\begin{aligned}
\mathbb{E}_{D,\epsilon} [(y - \hat{y})^2] &= \mathbb{E}_{D,\epsilon} [(g(x) + \epsilon - f(x; D))^2] \\
&= \mathbb{E}_{D,\epsilon} [(g(x) - f(x; D))^2 + 2\epsilon(g(x) - f(x; D)) + \epsilon^2] \\
&= \mathbb{E}_D [(g(x) - f(x; D))^2] + 2\mathbb{E}_{D,\epsilon} [\epsilon(g(x) - f(x; D))] + \mathbb{E}_\epsilon [\epsilon^2] \\
&= \mathbb{E}_D [(g(x) - f(x; D))^2] + \sigma^2 \quad (\text{ya que } \mathbb{E}[\epsilon] = 0 \text{ y } \mathbb{E}[\epsilon^2] = \sigma^2) \\
&= \mathbb{E}_D [(g(x) - \mathbb{E}_D[f(x; D)] + \mathbb{E}_D[f(x; D)] - f(x; D))^2] + \sigma^2 \\
&= \mathbb{E}_D [(g(x) - \mathbb{E}_D[f(x; D)])^2 + (\mathbb{E}_D[f(x; D)] - f(x; D))^2 \\
&\quad + 2(g(x) - \mathbb{E}_D[f(x; D)])(\mathbb{E}_D[f(x; D)] - f(x; D))] + \sigma^2 \\
&= \mathbb{E}_D [(g(x) - \mathbb{E}_D[f(x; D)])^2] + \mathbb{E}_D [(\mathbb{E}_D[f(x; D)] - f(x; D))^2] \\
&\quad + 2\mathbb{E}_D [(g(x) - \mathbb{E}_D[f(x; D)])(\mathbb{E}_D[f(x; D)] - f(x; D))] + \sigma^2 \\
&= (g(x) - \mathbb{E}_D[f(x; D)])^2 + \mathbb{E}_D [(f(x; D) - \mathbb{E}_D[f(x; D)])^2] + \sigma^2 \\
&= \underbrace{(g(x) - \mathbb{E}_D[f(x; D)])^2}_{\text{bias}^2} + \underbrace{\mathbb{E}_D [(f(x; D) - \mathbb{E}_D[f(x; D)])^2]}_{\text{varianza}} + \underbrace{\sigma^2}_{\text{error irreducible}}
\end{aligned}$$