CSCI 145 Problem 2 Solution: The Optimal Control Variate

Solution

Approach

The adjusted output has a variance that is quadratic in c. We find its minimum by differentiating with respect to c.

1. Variance and Optimal Control

For any sample $\mathbf{X}_i$,

$$\begin{aligned}
&\operatorname{Var}\left(F(\mathbf{X}_i) - c\left(G(\mathbf{X}_i) - \mathbb{E}[G(\mathbf{X})]\right)\right) \\
&= \operatorname{Var}\left(F(\mathbf{X}_i) - cG(\mathbf{X}_i)\right) \\
&= \operatorname{Var}\left(F(\mathbf{X}_i)\right) - 2c\operatorname{Cov}\left(F(\mathbf{X}_i), G(\mathbf{X}_i)\right) + c^2\operatorname{Var}\left(G(\mathbf{X}_i)\right) \\
&= \operatorname{Var}(F(\mathbf{X})) - 2c\operatorname{Cov}(F(\mathbf{X}), G(\mathbf{X})) + c^2\operatorname{Var}(G(\mathbf{X})).
\end{aligned}$$

The first equality removes the constant $c\mathbb{E}[G(\mathbf{X})]$, which does not affect variance. The last equality uses that every $\mathbf{X}_i$ has the same distribution as $\mathbf{X}$. Because the n adjusted outputs are independent and identically distributed,

$$\operatorname{Var}(\widehat{\mu}_{n,c}) = \frac{1}{n}\left(\operatorname{Var}(F(\mathbf{X})) - 2c\operatorname{Cov}(F(\mathbf{X}), G(\mathbf{X})) + c^2\operatorname{Var}(G(\mathbf{X}))\right).$$

Differentiate this variance with respect to c:

$$\frac{d}{dc}\operatorname{Var}(\widehat{\mu}_{n,c}) = \frac{1}{n}\left(-2\operatorname{Cov}(F(\mathbf{X}), G(\mathbf{X})) + 2c\operatorname{Var}(G(\mathbf{X}))\right).$$

Setting the derivative equal to zero gives

$$\begin{aligned}
0 &= -2\operatorname{Cov}(F(\mathbf{X}), G(\mathbf{X})) + 2c\operatorname{Var}(G(\mathbf{X})), \\
c^\star &= \frac{\operatorname{Cov}(F(\mathbf{X}), G(\mathbf{X}))}{\operatorname{Var}(G(\mathbf{X}))}.
\end{aligned}$$

The second derivative is positive:

$$\frac{d^2}{dc^2}\operatorname{Var}(\widehat{\mu}_{n,c}) = \frac{2\operatorname{Var}(G(\mathbf{X}))}{n} > 0.$$

Thus $c^\star$ is the unique global minimizer.

2. The Variance Reduction

Substituting it into the variance gives

$$\begin{aligned}\operatorname{Var}(\widehat{\mu}_{n,c^\star}) &= \frac{1}{n}\left(\operatorname{Var}(F(\mathbf{X})) - \frac{2\operatorname{Cov}(F(\mathbf{X}),G(\mathbf{X}))^2}{\operatorname{Var}(G(\mathbf{X}))} + \frac{\operatorname{Cov}(F(\mathbf{X}),G(\mathbf{X}))^2}{\operatorname{Var}(G(\mathbf{X}))}\right)\\ &= \frac{\operatorname{Var}(F(\mathbf{X}))}{n}\left(1 - \frac{\operatorname{Cov}(F(\mathbf{X}),G(\mathbf{X}))^2}{\operatorname{Var}(F(\mathbf{X}))\operatorname{Var}(G(\mathbf{X}))}\right).\end{aligned}$$