MA 323

Lecture # 10

Antithetic Variables:

The method of "antithetic variates" attemps to reduce variance by introducing negative dependence between pairs of replications. The most broadly applicable form for this is motivated from the fact that if $U \sim \mathcal{U}[0,1]$, then $(1-U) \sim \mathcal{U}[0,1]$, too.

Accordingly, if a path is generated using the inputs $U_1, U_2, \ldots U_n$, then we can generate a second path using $1-U_1, 1-U_2, \ldots, 1-U_n$, without changing the law of the simulated process. The variables U_i and $(1-U_i)$ form an antithetic pair in the sense that a large value of one is accompanied by a small value of the other.

Observation: An unusually large or small output computed from the first path may be balanced by the value computed from the antithetic path, resulting in reduction of variance.

These observations extend to other distributions, through inverse transform method: $F^{-1}(U)$ and $F^{-1}(1-U)$ both have distribution

F, but are antithetic to each other, because F^{-1} is monotone.

For a distribution symmetric about the origin, $F^{-1}(1-u)$ and $F^{-1}(u)$ have the same magnitudes, but opposite signs. In particular, a simulation driven by $Z_1, Z_2, \ldots$ of i.i.d. $N(0,1)$ variables along with $-Z_1, -Z_2, \ldots$ of i.i.d $N(0,1)$ variables, is an example of simulation driven by antithetic variates.

More specifically, our objective is to estimate an expectation $E[Y]$, and that using some implementation of antithetic sampling, produces a sequence of pairs of observations $(Y_1, \tilde{Y}_1), (Y_2, \tilde{Y}_2), \ldots, (Y_n, \tilde{Y}_n)$. Some key characteristic of antithetic variates method are the following:

1) The pairs $(Y_1, \tilde{Y}_1), (Y_2, \tilde{Y}_2), \ldots (Y_n, \tilde{Y}_n)$ are i.i.d

② For each i, Y_i and $\tilde{Y}_i$ have the same distribution, though ordinarily they are not independent.

<u>Notation</u> Y: Random variable with the common distribution of Y_i and $\tilde{Y}_i$.

The antithetic variates estimator is simply the average of all $2n$ observations,

$$\hat{Y}_{AV} = \frac{1}{2n}\left(\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} \tilde{Y}_i\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_i + \tilde{Y}_i}{2}\right).$$

$\hat{Y}_{AV}$ is the sample mean of the "n" independent observations:

$$\frac{Y_1+\tilde{Y}_1}{2}, \quad \frac{Y_2+\tilde{Y}_2}{2}, \ldots\ldots, \quad \frac{Y_n+\tilde{Y}_n}{2}.$$

Applying Central Limit Theorem, we get,

$$\frac{\hat{Y}_{AV} - E[Y]}{\sigma_{AV}/\sqrt{n}} \sim N(0,1)$$

with

$$\sigma_{AV}^2 = Var\left[\frac{Y_i + \tilde{Y}_i}{2}\right]$$

This limit in distribution continues to hold if we replace σ_{AV} with s_{AV} (the sample standard deviation of the n - values $\frac{Y_i + \tilde{Y}_i}{2}$, $i = 1, 2, \ldots, n$) .

Then a $(1-\delta)$ confidence interval is of the form :

$$\hat{Y}_{AV} \pm z_{\delta/2} \frac{s_{AV}}{\sqrt{n}}, \quad \text{where}$$

$$1 - \Phi(z_{\delta/2}) = \frac{\delta}{2}.$$

<u>Assumption:</u> Computational effort to generate $(Y_i, \tilde{Y}_i)$ is approximately twice the effort required to generate Y_i. The computational cost of generating inputs is a small fraction of the total cost of simulating Y_i. Under this

assumption, the effort required to compute $\hat{Y}_{AV}$ is approximately that required to compute the sample mean of $2n$ independent replications. Accordingly, it is meaningful to compare the variances of these two estimates. Usage of antithetics results in variance reduction if

$$\mathrm{Var}\left[\hat{Y}_{AV}\right] < \mathrm{Var}\left[\frac{1}{2n}\sum_{i=1}^{2n} Y_i\right];$$

i.e., if, $\mathrm{Var}[Y_i + \tilde{Y}_i] < 2\,\mathrm{Var}[Y_i]$.

Now $\mathrm{Var}[Y_i + \tilde{Y}_i] = \mathrm{Var}[Y_i] + \mathrm{Var}[\tilde{Y}_i] + 2\,\mathrm{Cov}(Y_i, \tilde{Y}_i)$

$= 2\,\mathrm{Var}[Y_i] + 2\,\mathrm{Cov}(Y_i, \tilde{Y}_i)$

$< 2\,\mathrm{Var}[Y_i]$

(Here, $\mathrm{Var}[Y_i] = \mathrm{Var}[\tilde{Y}_i]$, since Y_i and $\tilde{Y}_i$ have the same distribution)

Thus variance reduction using antithetic sampling is achieved if

$$\operatorname{Cov}\left[Y_i, \tilde{Y}_i\right] < 0 .$$

Consequence: Negative dependence in the inputs produces negative correlation between the outputs of paired replications.

* Condition ensuring the monotonicity of the mapping from inputs to outputs.

Suppose that the inputs to a simulation are independent random variables $X_1, X_2, \ldots, X_m$.

Suppose that Y is an increasing function of these inputs. Then,

$$E[Y\tilde{Y}] \leq E[Y]E[\tilde{Y}]$$

The requirement that the simulation map inputs to outputs monotonically, is rarely satisfied exactly, but provides some qualitative insight into the scope of the method.

Variance Decomposition: Antithetic variates eliminate the variance due to the antisymmetric part of an integrand, in the sense presented below.

Accordingly, we focus our attention to the case of standard normal inputs. But the observations are applicable to any other distributions symmetric about the origin and apply with minor modifications to uniformly distributed inputs.

Let $Y = f(z)$, $z = (z_1, z_2, \ldots, z_d) \sim N(0, I)$

Define ① Symmetric part of f: $f_0(z) = \dfrac{f(z) + f(-z)}{2}$

(2) Antisymmetric part of f: $f_1(z) = \dfrac{f(z) - f(-z)}{2}$

Clearly $f = f_0 + f_1$. Also,

$$E\left[f_0(z)\, f_1(z) \right] = \frac{1}{4} E\left[f^2(z) - f^2(-z) \right] = 0 = E\left[f_0(z) \right] E\left[f_1(z) \right].$$

Therefore $$\mathrm{Var}\left[f(z) \right] = \mathrm{Var}\left[f_0(z) \right] + \mathrm{Var}\left[f_1(z) \right]$$

Thus antithetic sampling eliminates all variances if f

is antisymmetric $(f = f_1)$ and it eliminates no variance
if f is symmetric $(f = f_0)$.