Chapter 3

Derivatives

In this chapter we approach the main problem of finding the minimizer of

$$F(x) = \int H(x, \omega)\, d\mu_x(\omega)$$

by discussing various notions of differentiability of parameter integrals.

The theoretical concepts presented in this chapter are the background for the simulation methods of the subsequent chapter 4. There we discuss some procedures to estimate not only the objective function $F(x)$ by a Monte-Carlo estimate $\widehat{F}(x)$, but also to estimate the gradient $\nabla F(x)$ by some $\widehat{\nabla F}(x)$.

It is important to stress that the representation of an objective function $F(x)$, which is of the form of a parameter integral $F(x) = \int H(x, u)\, d\mu_x$ is not unique. In particular, distinguishing between parameter integrals with *parameterized integrands* (here x is a *structural parameter*):

$$\int H(x, w)\, d\mu(w) \tag{I}$$

and parameter integrals with *parameterized integrators* (here x is a *distributional parameter*):

$$\int H(w)\, d\mu_x(w) \tag{II}$$

it is possible to represent a problem alternatively in form (I) and form (II).

3.1 Example. A retention basin for a river has to be designed in order to prevent floods. The construction costs are $c_1 x$, where x is the capacity of the basin. If the capacity is insufficient, a flood occurs, causing costs of c_2. Let ξ denote the random amount of inflowing water. The problem is

$$\left\| \begin{array}{l} \text{Minimize } c_1 x + c_2 \mathbb{E}(\mathbb{1}_{\{\xi > x\}}) \\ x \geq 0 \end{array} \right. \tag{3.1}$$

G. C. Pflug, *Optimization of Stochastic Models*

Since the linear part $c_1 x$ is easy to deal with, let us concentrate on the function $F_1(x) := \mathbb{E}(\mathbb{1}_{\{\xi > x\}})$. Denoting the distribution of ξ by μ, we may write

$$F_1(x) = \int H(x, w)\, d\mu(w), \qquad \text{(form (I))}$$

where

$$H(x, w) = \mathbb{1}_{\{w > x\}}.$$

On the other hand, we may write $\mathbb{1}_{\{\xi > x\}} = \mathbb{1}_{\{\xi - x > 0\}}$ and denoting the distribution of $\xi - x$ by μ_x (a translation family), we get

$$F_1(x) = \int H(w)\, d\mu_x(w), \qquad \text{(form (II))}$$

where

$$H(w) = \mathbb{1}_{\{w > 0\}}.$$

Further, suppose that there is a measure ν dominating all μ_x. Then, denoting the densities by $\psi_x(w)$, i.e. $d\mu_x = \psi_x d\nu$, one may write by change of measure

$$F_1(x) = \int H_1(x, w)\, d\nu(w), \qquad \text{(form (I))}$$

where

$$H_1(x, w) = \mathbb{1}_{\{w > 0\}} \psi_x(w).$$

Thus, we may go back and forth between representations (I) and (II). Rubinstein calls the switch from (I) to (II) push-out and the switch from (II) to (I) push-in (see Rubinstein (1992)).

Even if both the integrand H and the integrator (the measure) μ depends on x, one may simplify the situation by assuming that either H or μ depends on the decision x, but not both. This follows from the general rules of differentiation:

$$\begin{aligned} \frac{\partial}{\partial x} \int H(x, \omega)\, d\mu_x(\omega) &= \frac{\partial}{\partial y} \int H(y, \omega)\, d\mu_x(\omega) \,\Big|_{y=x} \\ &\quad + \frac{\partial}{\partial y} \int H(x, \omega)\, d\mu_y(\omega) \,\Big|_{y=x}. \end{aligned}$$

We will therefore study the differentiability properties of

$$x \mapsto \int H(x, \omega)\, d\mu(\omega)$$

– the derivatives of random processes – and of

$$x \mapsto \int H(\omega)\, d\mu_x(\omega)$$

– the derivatives of probability measures – separately in section 3.1 and section 3.2.