Differentiating Rollout Policies

1 Differentiating Argmax

Consider the maximization problem

$$\mathbf{x}^* := h(y) = \arg\max_x f(\mathbf{x}, y).$$

We want to find the derivative of $h(y)$ with respect to y. Assume that both $\mathbf{x}$ and y are in $\mathbb{R}$, so that $h(y) : \mathbb{R} \to \mathbb{R}$. Further assume that the argmax is also a local max, e.g.

$$\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^*, y) = 0,$$
$$\frac{\partial^2 f}{\partial \mathbf{x}^2}(\mathbf{x}^*, y) < 0.$$

We rewrite the first condition as:

$$\frac{\partial f}{\partial \mathbf{x}}(h(y), y) = 0.$$

Then we take the partial on both sides with respect to y. By the chain rule:

$$\frac{\partial^2 f}{\partial \mathbf{x} \partial y}(h(y), y) + \frac{dh}{dy} * \frac{\partial^2 f}{\partial \mathbf{x}^2}(h(y), y) = 0.$$

Solving for $\frac{\partial h}{\partial y}$:

$$\frac{dh}{dy} = -\frac{\partial^2 f}{\partial \mathbf{x}^2}(h(y), y)^{-1} \frac{\partial^2 f}{\partial \mathbf{x} \partial y}(h(y), y) = -\frac{\partial^2 f}{\partial \mathbf{x}^2}(\mathbf{x}^*, y)^{-1} \frac{\partial^2 f}{\partial \mathbf{x} \partial y}(\mathbf{x}^*, y).$$

If $\mathbf{x} \in \mathbb{R}^d$, then $h(y) : \mathbb{R} \to \mathbb{R}^d$ and its partial derivative (vector-valued) is:

$$\frac{dh}{dy} = H_{\mathbf{x}^*}^{-1} d_{\mathbf{x}^*},$$

where $H_{\mathbf{x}^*}$ is the Hessian of f evaluated at $\mathbf{x}^*$ ($H_{\mathbf{x}^*}$ is invertible because it's negative-definite), and the i-th entry of $d_{\mathbf{x}^*}$ is the partial $\frac{\partial^2 f}{\partial \mathbf{x}_i \partial y}(\mathbf{x}^*, y)$.

1.1 When the objective is EI

When the objective is EI, then H is the Hessian of EI with respect to $\mathbf{x}$. $\frac{\partial^2 f}{\partial \mathbf{x}_i \, \partial y}(\mathbf{x}^*, y)$ is a little trickier. EI is:

$$\mathrm{EI}(\mathbf{x}, y) = (y^* - \mu(\mathbf{x}, y))\Phi(\mathbf{x}, y) + \sigma(\mathbf{x})\phi(\mathbf{x}, y).$$

We have indicated a dependence on y where relevant. $\mu(\mathbf{x}, y)$ is the posterior mean, $\sigma(\mathbf{x})$ is the posterior standard deviation, $\phi(\mathbf{x}, y)$ is the posterior pdf, and $\Phi(\mathbf{x}, y)$ is the posterior CDF. We will need to differentiate μ, ϕ, and Φ w.r.t. y. I haven't worked this part out (should be something that involves a Schur complement).

2 Differentiating Rollout of EI

Let's consider the case of $h = 2$ for now (horizon 2). We call this EI 2:

$$\mathrm{EI}\,2(\mathbf{x}) := \mathbb{E}[(y^* - y_2(\mathbf{x})].$$

where $y_2(\mathbf{x})$ is a sampled value from the posterior GP after a step of BO has already performed. The main challenge behind the gradient of EI $2(\mathbf{x})$ is the gradient of $y_2(\mathbf{x})$. The sequence of computation to sample $y_2(\mathbf{x})$ is:

$$\mathbf{x} \to y_1(\mathbf{x}) \to \mathbf{x}_2 = \arg\max_{\mathbf{x}'} \mathrm{EI}(\mathbf{x}', y_1(\mathbf{x})) \to y_2(\mathbf{x}_2).$$

We denote $\arg\max_{\mathbf{x}'} \mathrm{EI}(\mathbf{x}', y_1(\mathbf{x}))$ as $h(y_1(\mathbf{x}))$. We thus rewrite y_2 implicitly:

$$y_2(\mathbf{x}) := y_2(h(y_1(\mathbf{x}))).$$

We differentiate:

$$\nabla y_2(\mathbf{x}) = \nabla y_2(h(y_1(\mathbf{x}))) * \frac{dh}{dy}(y_1(\mathbf{x})) * \nabla y_1(\mathbf{x}).$$

Note the multiplication is element-wise. We can generalize this to longer horizons. For example, for $h = 3$:

$$y_3(\mathbf{x}) := y_3(h_2(y_2(h_1(y_1(\mathbf{x}))))),$$

and differentiation yields:

$$\begin{aligned}\nabla y_3(\mathbf{x}) = &\nabla y_3(h_2(y_2(h_1(y_1(\mathbf{x})))))* \\ &\frac{dh_2}{dy}(y_2(h_1(y_1(\mathbf{x}))))* \\ &\nabla y_2(h_1(y_1(\mathbf{x}))))* \\ &\frac{dh_1}{dy}(y_1(\mathbf{x}))* \\ &\nabla y_1(\mathbf{x}).\end{aligned}$$