Bayesian Optimization

Exercise 1: Expected Improvement

We start with

$$a_{\mathrm{EI}}(\mathbf{x}) = \mathbb{E}_y(\max\{f_{\min} - y, 0\}) = \int_{-\infty}^{\infty} \max\{f_{\min} - y, 0\} p(y) dy.$$

Observe that

$$\max\{f_{\min} - y, 0\} = \begin{cases} f_{\min} - y, & \text{if } y < f_{\min}, \\ 0, & \text{otherwise.} \end{cases}$$

All contributions for $y \geq f_{\min}$ are zero. Therefore, we can additively decompose the integral and it simplifies to

$$a_{\mathrm{EI}}(\mathbf{x}) = \int_{-\infty}^{f_{\min}} (f_{\min} - y)\, p(y) dy.$$

$$\begin{aligned}
\alpha_{\mathrm{EI}}(\mathbf{x}) &= \int_{-\infty}^{f_{\min}} (f_{\min} - y)\, p(y) dy \\
&= \int_{-\infty}^{f_{\min}} (f_{\min} - y)\, \frac{1}{\sqrt{2\pi \hat{s}(\mathbf{x})^2}} \exp\left(-\frac{\left(y - \hat{f}(\mathbf{x})\right)^2}{2\hat{s}(\mathbf{x})^2}\right) dy \\
&= \int_{-\infty}^{z} \left(f_{\min} - \hat{f}(\mathbf{x}) - u\hat{s}(\mathbf{x})\right) \frac{1}{\sqrt{2\pi \hat{s}(\mathbf{x})^2}} \exp\left(-\frac{u^2}{2}\right) \hat{s}(\mathbf{x}) du \quad \left(\text{Def. } u := \frac{y - \hat{f}(\mathbf{x})}{\hat{s}(\mathbf{x})}, \frac{du}{dy} = \frac{1}{\hat{s}(\mathbf{x})}, z := \frac{f_{\min} - \hat{f}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) \\
&= \int_{-\infty}^{z} \left(f_{\min} - \hat{f}(\mathbf{x}) - u\hat{s}(\mathbf{x})\right) \phi(u) du \\
&= \int_{-\infty}^{z} \left(f_{\min} - \hat{f}(\mathbf{x})\right) \phi(u) du - \int_{-\infty}^{z} (u\hat{s}(\mathbf{x}))\, \phi(u) du
\end{aligned}$$

Note that

$$\Phi(z) = \int_{-\infty}^{z} \phi(u) du$$

by definition.

Therefore, regarding the first integral:

$$\int_{-\infty}^{z} \left(f_{\min} - \hat{f}(\mathbf{x})\right) \phi(u) du = \left(f_{\min} - \hat{f}(\mathbf{x})\right) \Phi(z) = z\hat{s}(\mathbf{x})\Phi(z).$$

Regarding the second integral we use the identity

$$\int_{-\infty}^{z} u\phi(u) du = -\phi(z).$$

Putting both together we obtain:

$$\begin{aligned}
\alpha_{\mathrm{EI}}(\mathbf{x}) &= z\hat{s}(\mathbf{x})\Phi(z) - \hat{s}(\mathbf{x})(-\phi(z)) \\
&= z\hat{s}(\mathbf{x})\Phi(z) + \hat{s}(\mathbf{x})\phi(z) \\
&= \left(f_{\min} - \hat{f}(\mathbf{x})\right) \Phi\left(\frac{f_{\min} - \hat{f}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right) + \hat{s}(\mathbf{x})\phi\left(\frac{f_{\min} - \hat{f}(\mathbf{x})}{\hat{s}(\mathbf{x})}\right).
\end{aligned}$$

Exercise 2: BO Loop

(a) Let $\mathcal{D}$ be the initial design consisting of $\{(x^{[1]}, y^{[1]}), \dots, (x^{[4]}, y^{[4]})\}$. Set t to 4.
While $t < 10$:

(i) Fit surrogate model on $\mathcal{D}$.

(ii) Optimize the Expected Improvement $a_{\mathrm{EI}}(x)$ to obtain a new point $x^{[t+1]} := \arg\max_{x \in [0,1]} a_{\mathrm{EI}}(x)$.

(iii) Evaluate $x^{[t+1]}$ and update design data $\mathcal{D} = \mathcal{D} \cup \{(x^{[t+1]}, f(x^{[t+1]}))\}$.

(iv) Set t to $t + 1$.

Return x that minimizes $f(x)$ in $\mathcal{D}$: $\arg\min_{(x,y) \in \mathcal{D}} y$.

(b)
```r
library(DiceKriging)
set.seed(0308)
f = function(x) 2*x * sin(14*x)
initial_x = runif(4, min = 0, max = 1)
initial_y = f(initial_x)
design = data.frame(x = initial_x, y = initial_y)
t = 4

ei = function(x, current_fmin, current_gp) {
  gp_prediction = predict(current_gp, newdata = data.frame(x = x), type="SK")
  gp_mean = gp_prediction$mean
  gp_sd = gp_prediction$sd
  diff = (current_fmin - gp_mean)
  z = diff / gp_sd
  diff * pnorm(z) + gp_sd * dnorm(z)
}


while (t < 10) {
  gp = km(design = design[, 1L, drop = FALSE], response = design[, 2L],
        covtype = "gauss", nugget = 1e-8)
  fmin = min(design$y)
  x_new = optimize(f = ei, interval = c(0, 1), maximum = TRUE,
        current_fmin = fmin, current_gp = gp)$maximum
  design = rbind(design, data.frame(x = x_new, y = f(x_new)))
  t = t + 1
}

design[which.min(design$y), ]
```