Homework 1

Stats 102B Lec 1 and 2

Spring 2024

General Guidelines

Show all your work, including any and all relevant code and output. Any and all collaboration must adhere to **Level 1** collaboration described in the Stats 102B Collaboration Policy.

Please submit your homework as a single file, in PDF format only. Name your assignment file with the convention `123456789_stats102b_hw0.pdf`, where `123456789` is replaced with your UID and `hw0` is updated to the actual homework number. Include your first and last name and UID in your assignment as well.

All R code is expected to follow the Tidyverse Style Guide: https://style.tidyverse.org/. If you scan or take a picture of any written work, please make sure the resolution is high enough that your work is clear and legible. Submissions with severe style or formatting issues may receive a penalty. Any submission that cannot be properly read will not be graded.

Question 1

As mentioned in lecture, the question of what it means to be the very best Pokémon trainer may vary from person to person. Each trainer may have a different set of goals in mind that contributes different amount to personal fulfillment. For the following members of the Pokémon universe, loosely write a cost function that they may seek to maximize subject to any possible constraints. The first one has been done as an example. (This is a fun introductory question. Don't sweat it too much as there is not one correct answer.)

- Ash Ketchum wants to collect every single gym badge (there are 8 of them) and defeat the Elite Four (four of them). Let g represent the number of gym badges he has obtained and E represent the number of elite four he has defeated.

$$c(g, E) = g + E,\ 0 \leq g \leq 8,\ 0 \leq E \leq 4$$

- Dawn wants to collect every single Pokémon contest ribbon r (there are five of them) and each ribbon brings her an even greater sense of personal fulfillment. However, she wants to spend time t (measured in hours) with her mother Joanna as well. It takes about 10 hours to win each ribbon, and she wants to spend more time with her mom than the total amount of time it took winning ribbons.
- Nando's goal is to collect both gym badges and ribbon, though ribbons bring him double the fulfillment of gym badges. However, the Pokémon league caps the total number of badges and ribbons to a maximum of 12.
- Come up with your own example of a cost function that you maximize/minimize as you walk to Stats 102B.

Question 2

Write a function `aem(g, n, N, LB = -1, UB = 1, method = "deterministic", min = TRUE)` that runs the approximately exhaustive method with the following parameters

- `g` is a function $g : \mathbb{R}^n \to \mathbb{R}$
- `n` is the dimension of the function's domain
- `N` is number of points to sample PER dimension
- `LB` is the lower bound of the search region
- `UB` is the upper bound of the search space. In two dimensions, the space should be a square centered at the origin (the corners of the search space are then (UB, UB), (LB, UB), (UB, LB), (LB, LB).
- `method` is a characters string default set to `"deterministic"` but can also accommodate `"stochastic"`
- `min` is a logical default set to `TRUE` to find the global minimum of `g`

The output should be a list object with the following components:

- `$index` should represent w^*, the global minimum/maximum
- `$val` should represent $g(w^*)$, the value of the function at the global minimum/maximum
- `$eval_matrix`, a ($N^n \times (n+1)$ for deterministic or $N \times (n+1)$ for stochastic) matrix containing the points generated in the search space, with the last column being the values of $g(w_i)$.
- `$time`, the time in seconds it took to run the algorithm

Now check your function on the following examples.

Note: Code should *never* be copy-pasted from PDFs. Certain characters become non-standard characters when embedded into a PDF. To make sure you and the grader have no issues when knitting your Rmd file, you should *always* retype code provided in PDFs.

```r
f <- function(v) {
  v[1] ^ 2 + v[2] ^ 2
}

test_1 <- aem(f, n = 2, N = 100)
test_1$index
test_1$val

set.seed(24601)

test_2 <- aem(f, n = 2, N = 100, method = "stochastic")
test_2$index
test_2$val
```

```r
h <- function(v) {
  exp(abs(v[1]^2 - 1) + abs(v[2]^3 + abs(v[3])))
}

test_1 <- aem(h, n = 3, N = 100, LB = -2, UB = 2)
test_1$index
test_1$val

set.seed(24601)
test_2 <- aem(h, n = 3, N = 100, LB = -2, UB = 2, method = "stochastic")
test_2$index
test_2$val
```

```r
j <- function(v) {
  log(abs(sin(v[1]) + cos(v[2])) + 1)
}
```

```r
set.seed(24601)
test_2 <- aem(j, n = 2, N = 100, method = "stochastic")
test_2$index
test_2$val

test_2 <- aem(j, n = 2, N = 100, method = "stochastic", min = FALSE)
test_2$index
test_2$val
```

Question 3

Recall from Stats 100A: The probability distribution function of an exponential distribution with parameter $\lambda > 0$ is defined by

$$f(x|\lambda) = \lambda e^{-\lambda x}, \ x \geq 0$$

a. Find the maximum likelihood estimator for λ. (If you have forgotten how to do this, ask Josh on Campuswire)
b. Consider the observed values 0.368, 0.714, 0.126, 0.006, 0.525, 0.36, 0.095, 0.212, 0.107, 0.058. Calculate the maximum likelihood estimate.
c. Using the stochastic version of the approximately exhaustive method on the log-likelihood of the exponential distribution with LB $= 0$, UB $= 5$ and $N = 1000$, calculate the maximum likelihood estimate. Are the two estimates comparable?

Question 4

Consider the function

$$f(\vec{x}) = \sum_{i=1}^{n} x_i^2$$

where x_i is the ith component of $\vec{x} \in \mathbb{R}^n$. Illustrate the curse of dimensionality by running the stochastic version of the approximately exhaustive method for $n = 1, 2, 3, 4$ with $\text{LB} = -1, \text{UB} = 1, N = 30$. Include a plot of the time taken over n.

Question 5

Consider the random search, coordinate search, and coordinate descent algorithms.

a. Explain how the one-dimensional versions of each of the algorithms are essentially identical.
b. Write a function `zom(g, alpha, w0, K)` that runs the one-dimensional version of the algorithms above with the following parameters:

   - `g` is a function $g : \mathbb{R} \to \mathbb{R}$
   - `alpha` is the step size
   - `w0` is the initial point
   - `K` is the number of iterations

   The output should be a list object with the following components:

   - `$index` should represent w^*, the global minimum/maximum
   - `$val` should represent $g(w^*)$, the value of the function at the global minimum/maximum

c. Verify your algorithm by implementing the exercise in Chapter 2 slide 50. Be sure to show your output.
d. Verify your algorithm by optimizing the log-likelihood in 3c. Use $\alpha = 0.01, w^0 = 1, \ K = 1000$.
e. Play around with different values of α and K for the previous question (5d). A good starting point is $\alpha \in \{1, 0.1, 0.01, 0.001\}$ and $K \in \{10, 100, 1000, 10000\}$. What can you conclude?