

4강 연습문제

1. 이변수 함수

$$f(x, y) = xy$$

에 대하여 learning rate $\eta = 1$ 과 계수 $\beta_1 = \beta_2 = 1/2$ 로 Adam을 적용하려 한다.

- (i) 초기 위치 $\mathbf{x}_0 = (1, 2)$ 에서 출발하여 두 발자국 걸어갈때, $\mathbf{x}_1, \mathbf{x}_2$ 를 구하시오.
- (ii) 코드로 계산하시오.

2. 이변수 함수

$$f(x, y) = x^2 + xy$$

에 대하여 learning rate $\eta = 1$ 과 계수 $\beta_1 = \beta_2 = 1/2$ 로 Adam을 적용하려 한다.

- (i) 초기 위치 $\mathbf{x}_0 = (1, 1)$ 에서 출발하여 두 발자국 걸어갈때, $\mathbf{x}_1, \mathbf{x}_2$ 를 구하시오.
- (ii) 코드로 계산하시오.

3. 밑줄 친 부분을 수식으로 바꾸고 이로부터 Adam의 점화식을 이끌어 내시오.

```
class Adam:
    def __init__(self, lr=0.001, beta1=0.9, beta2=0.999):
        self.lr = lr
        self.beta1 = beta1
        self.beta2 = beta2
        self.iter = 0
        self.m = None
        self.v = None
    def update(self, params, grads):
        if self.m is None:
            self.m, self.v = {}, {}
            for key, val in params.items():
                self.m[key] = np.zeros_like(val)
                self.v[key] = np.zeros_like(val)
        self.iter += 1
        lr_t = self.lr * np.sqrt(1.0 - self.beta2**self.iter)
                / (1.0 - self.beta1**self.iter)
        for key in params.keys():
            self.m[key] += (1 - self.beta1) * (grads[key] - self.m[key])
            self.v[key] += (1 - self.beta2) * (grads[key]**2 - self.v[key])
            params[key] -= lr_t * self.m[key] / (np.sqrt(self.v[key]) + 1e-7)
```