CS4400 Deep Reinforcement Learning
Wendelin Böhmer `<j.w.bohmer@tudelft.nl>`

Exercise Sheet 1
voluntary exercises

Math and machine learning primer

Voluntary exercises

The following exercises do not have to be submitted as homework, but might be helpful to practice the required math and prepare for the exam. Some questions are from old exams and contain the used rubrik. You do not have to submit these questions and will not receive points for them.

E1.1: Taylor expansion **(voluntary)**

For the function $\sqrt{1+x}$, write down the Taylor series around $x_0 = 0$ up to 3rd order.

Solution Approximating $f(x)$ via Taylor expansion at x_0:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

i.e. for an expansion around $x_0 = 0$

$$f(x) \quad\approx\quad f(0) \quad+\quad f'(0)x \quad+\quad \frac{1}{2}f''(0)x^2 \quad+\quad \frac{1}{6}f'''(0)x^3 \quad+\quad \mathcal{O}(x^4)$$

with

$$\begin{aligned} f'(x) &= \frac{1}{2}(x+1)^{-\frac{1}{2}} && \to f'(0) = 1/2 \\ f''(x) &= -\frac{1}{4}(x+1)^{-\frac{3}{2}} && \to f''(0) = -1/4 \\ f'''(x) &= \frac{3}{8}(x+1)^{-\frac{5}{2}} && \to f'''(0) = 3/8 \end{aligned}$$

the Taylor expansion of $\sqrt{1+x} = (1+x)^{\frac{1}{2}}$ around $x_0 = 0$ reads

$$\sqrt{1+x} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + ...$$

E1.2: Critical points **(voluntary)**

Consider the two functions

$$\begin{aligned} f(x,y) &:= c + x^2 + y^2 \\ g(x,y) &:= c + x^2 - y^2, \end{aligned}$$

where $c \in \mathbb{R}$ is a constant.

(a) Show that $\boldsymbol{a} = (0, 0)$ is a critical point of both functions.

(b) Check for f and for g whether $\boldsymbol{a}$ is a minimum, maximum, or saddlepoint using the Hessian matrix.

Hint: A matrix is positive (negative) definite if and only if all its eigenvalues are positive (negative).

Solution

$$\nabla f(\boldsymbol{a}) = (2x, 2y)\Big|_{(x,y)=\boldsymbol{a}} = (0, 0)$$

and

$$\nabla g(\boldsymbol{a}) = (2x, -2y)\Big|_{(x,y)=\boldsymbol{a}} = (0, 0)$$

$\Longrightarrow$ Necessary condition of extrema (vanishing gradient) is fulfilled at $\boldsymbol{a}$.

Checking for extrema:

Minimum if Hessian matrix $\mathbf{H}$ is positive definite (all eigenvalues >0)

Maximum if $\mathbf{H}$ is negative definite (all eigenvalues <0)

$\rightarrow$ characteristic polyomial i.e. $\det[\mathbf{H} - \lambda\mathbf{I}]$:

$$(\mathbf{H}_f)(\boldsymbol{a}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \Rightarrow (2-\lambda)^2 \overset{!}{=} 0$$

i.e. all eigenvalues (2&2) are real, positive $\Rightarrow$ $\mathbf{H}_f$ is pos. definite. Thus, $\boldsymbol{a}$ is a minimum of f

$$(\mathbf{H}_g)(\boldsymbol{a}) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \quad \Rightarrow (2-\lambda)(-2-\lambda) \overset{!}{=} 0$$

positive and negative Eigenvalues (2&-2) $\Rightarrow$ $\mathbf{H}_g$ is neither positive nor negative definite. Therefore $\boldsymbol{a}$ is a saddlepoint but no extremum of g.

E1.3: Distributions and expected values **(voluntary)**

Let $x \in \mathbb{R}$ be a random variable with probability density $p : \mathbb{R} \rightarrow \mathbb{R}$ with:

$$p(x) = \begin{cases} c \cdot \sin(x), & x \in [0, \pi] \\ 0, & \text{elsewhere} \end{cases}$$

(a) Determine the parameter $c \in \mathbb{R}$ such that $p(x)$ is indeed a probability density.

(b) Determine the expected value $\mu := \mathbb{E}_p[x]$

(c) Determine the variance of x, $\mathbb{E}_p[(x-\mu)^2]$.

Solution

(a) For p being a probability density it is required that (i) $p(x) \geq 0 \; \forall x \in \mathbb{R}$ which is fullfilled here. Furthermore, (ii) p must be normalized appropriately:

$$\int_{\mathbb{R}} p(x)dx \quad = \quad 1$$

Therefore, we get for the unknown constant c:

$$c\int_0^\pi \sin(x)dx \quad = \quad c[-\cos(x)]\Big|_0^\pi \quad = \quad 2c \quad \overset{!}{=} \quad 1 \quad \rightarrow \quad c = 1/2$$

(b) To calculate the expected value, we use integration by parts, i.e., for any functions f and g:

$$\int_a^b fg'dx \quad = \quad (fg)\Big|_a^b - \int_a^b f'gdx$$

$$\mu := \mathbb{E}_p[x] = \quad \tfrac{1}{2}\int_0^\pi x\sin(x)dx \quad = \quad -\tfrac{1}{2}x\cos(x)\Big|_0^\pi + \tfrac{1}{2}\underbrace{\int_0^\pi \cos(x)dx}_{=0} \quad = \pi/2$$

(c) To calculate the variance, we proceed in the same way

$$\mathbb{E}_p[x^2] = \tfrac{1}{2}\int_0^\pi x^2\sin(x)dx \quad = \underbrace{-\tfrac{1}{2}x^2\cos(x)\Big|_0^\pi}_{=\frac{\pi^2}{2}} \quad + \quad \underbrace{\tfrac{2}{2}\int_0^\pi x\cos(x)dx}_{=k}$$

with

$$k \quad = \quad x\sin(x)\Big|_0^\pi - \int_0^\pi \sin(x)dx \quad = \quad 0 + \cos(x)\Big|_0^\pi \quad = 0 - 2$$

and therefore

$$\mathbb{E}_p[x^2] \quad = \quad \tfrac{\pi^2}{2} - 2$$

yielding

$$\mathbb{E}_p[(x-\mu)^2] \quad = \quad \mathbb{E}_p[x^2] - \mu^2 \quad = \quad \tfrac{\pi^2}{2} - 2 - \tfrac{\pi^2}{4} \quad = \quad \tfrac{\pi^2}{4} - 2\,.$$

E1.4: Variance of the empirical mean (old exam question) (voluntary)

Prove that the variance of the empirical mean $f_n := \frac{1}{n}\sum_{i=1}^n x_i$, based on n samples $x_i \in \mathbb{R}$ drawn i.i.d. from the Gaussian distribution $\mathcal{N}(\mu,\sigma^2)$, is $\mathbb{V}[f_n] = \frac{\sigma^2}{n}$, *without* using the fact the variance of a sum of independent variables is the sum of the variables' variances.

<u>Solution</u> The major insights are that $\mathbb{E}[x_i] = \mu, \forall i$, $\mathbb{E}[x_i x_j] = \mathbb{E}[x_i]\mathbb{E}[x_j]$ if $i \neq j$ due to i.i.d. sampling and that $\mathbb{E}[(x_i-\mu)^2] = \sigma^2$.

$$\begin{aligned}\mathbb{V}[f_n] \quad &= \quad \mathbb{E}\Big[\big(\tfrac{1}{n}\sum_{i=1}^n x_i - \mu\big)^2\Big] \quad = \quad \tfrac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \mathbb{E}[(x_i-\mu)(x_j-\mu)] \\ &= \quad \tfrac{1}{n^2}\sum_{i\neq j}\underbrace{\mathbb{E}[(x_i-\mu)]}_{0}\underbrace{\mathbb{E}[(x_j-\mu)]}_{0} + \tfrac{1}{n^2}\sum_{i=1}^n\underbrace{\mathbb{E}[(x_i-\mu)^2]}_{\sigma^2} \quad = \quad \tfrac{\sigma^2}{n}\,.\end{aligned}$$

Rubrik:

- 1 point for the correct definition of variance $\mathbb{V}$
- 1 point for using $\mathbb{E}[x_i] = \mu$
- 1 point for the use of independent samples
- 1 point for the use of the definition of σ^2
- 1 point for putting it correctly together
- $-\frac{1}{2}$ points for minor mistakes (e.g. $\mathbb{E}[x_i x_j] = 0$ for $i \neq j$)
- but no point loss for forgetting little things like one or two ± mistakes

E1.5: Unbiased variance estimate **(voluntary)**

Let $\{x_i\}_{i=1}^n$ be a data set that is drawn i.i.d. from the Gaussian distribution $x_i \sim \mathcal{N}(\mu, \sigma^2)$. Let further $\hat\mu := \frac{1}{n}\sum_{i=1}^n x_i$ denote the empirical mean and $\hat\sigma^2 := \frac{1}{n}\sum_{i=1}^n (x_i - \hat\mu)^2$ the equivalent empirical variance. Prove analytically that $\hat\mu$ is unbiased, i.e. $\mathbb{E}[\hat\mu] = \mu$, and that $\hat\sigma^2$ is biased, i.e. $\mathbb{E}[\hat\sigma^2] \neq \sigma^2$.

Bonus-question: Can you derive an unbiased estimator for the empirical variance?

Hint: If x_i and x_j are drawn i.i.d. from $\mathcal{N}(\mu, \sigma^2)$, then holds $\forall i$:

$$\mathbb{E}[x_i] = \mu\,, \qquad \mathbb{E}[(x_i - \mu)^2] = \sigma^2 \qquad \text{and} \qquad \mathbb{E}[(x_i - \mu)(x_j - \mu)] = 0 \quad \text{if} \quad i \neq j\,.$$

<u>Solution</u> We prove that $\hat\mu$ is bias free simply by using its definition:

$$\mathbb{E}[\hat\mu] \quad = \quad \mathbb{E}\Big[\tfrac{1}{n}\sum_{i=1}^n x_i\Big] \quad = \quad \tfrac{1}{n}\sum_{i=1}^n \underbrace{\mathbb{E}[x_i]}_{\mu} \quad = \quad \mu\,.$$

Proving that $\hat\sigma^2$ is biased is more involved, as $\hat\sigma^2$ contains the empirical mean $\hat\mu$:

$$\begin{aligned}
\mathbb{E}[\hat\sigma^2] \quad &= \quad \tfrac{1}{n}\sum_{i=1}^n \mathbb{E}\Big[(x_i - \hat\mu)^2\Big] \quad = \quad \tfrac{1}{n}\sum_{i=1}^n \mathbb{E}[x_i^2] - 2\tfrac{1}{n}\sum_{i=1}^n \mathbb{E}[x_i \hat\mu] + \mathbb{E}[\hat\mu^2] \\
&= \quad \tfrac{1}{n}\sum_{i=1}^n \mathbb{E}[x_i^2] - 2\tfrac{1}{n}\sum_{i=1}^n \mathbb{E}\Big[x_i\, \tfrac{1}{n}\sum_{j=1}^n x_j\Big] + \mathbb{E}\Big[\tfrac{1}{n}\sum_{i=1}^n x_i\, \tfrac{1}{n}\sum_{j=1}^n x_j\Big] \\
&= \quad \tfrac{1}{n}\sum_{i=1}^n \mathbb{E}[x_i^2] - \tfrac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \mathbb{E}[x_i x_j] \underbrace{-\mu^2 + \mu^2}_{0} \\
&= \quad \tfrac{1}{n}\sum_{i=1}^n \underbrace{\mathbb{E}[(x_i - \mu)^2]}_{\sigma^2} - \tfrac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \underbrace{\mathbb{E}[(x_i - \mu)(x_j - \mu)]}_{\sigma^2 \text{ if } i=j \text{ else } 0} \\
&= \quad \sigma^2 - \tfrac{1}{n}\sigma^2 \quad = \quad \tfrac{n-1}{n}\sigma^2\,,
\end{aligned}$$

where we used $\mathbb{E}[(x_i - \mu)(x_j - \mu)] = \mathbb{E}[x_i x_j] - \mathbb{E}[x_i]\mu - \mathbb{E}[x_j]\mu + \mu^2 = \mathbb{E}[x_i x_j] - \mu^2$, because $\mathbb{E}[x_i] = \mu$.

Bonus-question: Note that $\hat\sigma^2$ would be unbiased if we would multiply it with $\frac{n}{n-1}$ and we can therefore define the unbiased empirical estimate of the variance as

$$\hat{\hat\sigma}^2 \quad := \quad \tfrac{1}{n-1}\sum_{i=1}^n (x_i - \hat\mu)^2\,.$$

E1.6: Maximum dice **(voluntary)**

This question is designed to practice the use of Kronecker-delta functions and become more familiar with (discrete) probabilities. You are given 3 dice, a D6, a D8 and a D10, where Dx refers to a x-sided fair dice, where each of the x sides is numbered uniquely 1 to x and rolled with the exact same probability.

(a) Prove analytically that the probability that the D6 is among the highest (including equal) numbers when all 3 dice are rolled together is roughly $\rho \approx 19\%$.

(b) Prove analytically that the probability that the D8 rolls among the highest is $\rho' \approx 38\%$.

(c) Prove analytically that the probability that the D10 rolls among the highest is $\rho'' \approx 58\%$.

Hint: You can solve the question however you want, but you are encouraged to use Kronecker-deltas, e.g. $\delta(i > 5)$ is 1 if $i > 5$ and 0 otherwise. You will find that this can simplify complex sums enormously. If you do so, you can use the equalities $\sum_{i=1}^n i \overset{(1)}{=} \frac{n^2+n}{2}$ and $\sum_{i=1}^n i^2 \overset{(2)}{=} \frac{n(n+1)(2n+1)}{6}$.

Bonus-question: Why don't the above numbers sum up to 1?

<u>Solution</u> The three dice are statistically independent and have the probability $p_x(i) = \frac{1}{x}$ of outcome $1 \leq i \leq x$. The probability of a Dx rolling higher or equal than a Dy is therefore:

$$p(i \geq j | i \sim p_x, j \sim p_y) \quad = \quad \frac{1}{xy}\sum_{i=1}^{x}\sum_{j=1}^{y}\delta(i \geq j)\,.$$

Note that if two conditions must be true, one can simply multiply the Kronecker-delta functions.

(a) The probability ρ of a D6 to roll higher than the D8 *and* the D10 is thus:

$$\begin{aligned}\rho &= p(i \geq j \wedge i \geq k | i \sim p_6, j \sim p_8, k \sim p_{10}) \quad = \quad \frac{1}{6\cdot8\cdot10}\sum_{i=1}^{6}\overbrace{\sum_{j=1}^{8}\delta(i \geq j)}^{i}\,\overbrace{\sum_{k=1}^{10}\delta(i \geq k)}^{i} \\ &= \frac{1}{480}\sum_{i=1}^{6} i^2 \quad \overset{(2)}{=} \quad \frac{1}{480}\frac{6(6+1)(12+1)}{6} \quad = \quad \frac{91}{480} \quad \approx \quad 19\%\,.\end{aligned}$$

(b) The major difference is that $\sum_{j=1}^6 \delta(i \geq j)$ cannot get larger than 6, even if $i > 6$.

$$\begin{aligned}\rho' &= p(i \geq j \wedge i \geq k | i \sim p_8, j \sim p_6, k \sim p_{10}) \quad = \quad \frac{1}{6\cdot8\cdot10}\sum_{i=1}^{8}\sum_{j=1}^{6}\sum_{k=1}^{10}\delta(i \geq j)\,\delta(i \geq k) \\ &= \frac{1}{6\cdot8\cdot10}\sum_{i=1}^{8}\sum_{j=1}^{6}\delta(i \geq j)\,\underbrace{\sum_{k=1}^{10}\delta(i \geq k)}_{i} \quad = \quad \frac{1}{6\cdot8\cdot10}\sum_{i=1}^{8} i\,\underbrace{\sum_{j=1}^{6}\delta(i \geq j)}_{\min(i,6)} \\ &= \frac{1}{6\cdot8\cdot10}\Big(\sum_{i=1}^{6} i^2 + \sum_{i=7}^{8} 6i\Big) \quad \overset{(2)}{=} \quad \frac{1}{6\cdot8\cdot10}\Big(\frac{6\cdot7\cdot13}{6} + 6(7+8)\Big) \quad = \quad \frac{91+90}{480} \quad \approx \quad 38\%\,.\end{aligned}$$

(c) Similarly for the D10:

$$\begin{aligned}\rho'' &= p(i \geq j \wedge i \geq k | i \sim p_{10}, j \sim p_6, k \sim p_8) \quad = \quad \frac{1}{6\cdot8\cdot10}\sum_{i=1}^{10}\overbrace{\sum_{j=1}^{6}\delta(i \geq j)}^{\min(i,6)}\,\overbrace{\sum_{k=1}^{8}\delta(i \geq k)}^{\min(i,8)} \\ &= \frac{1}{480}\Big(\sum_{i=1}^{6} i^2 + \sum_{i=7}^{8} 6i + \sum_{i=9}^{10} 6\cdot 8\Big) \quad = \quad \frac{91+90+96}{480} \quad \approx \quad 58\%\,.\end{aligned}$$

Bonus-question: Because conditions like $\delta(i \geq j)$ and $\delta(j \geq i)$ overlap in the case $\delta(i = j)$. To get a probability distribution over *disjunct* outcomes, one would have to consider the cases "D6 is *highest*, D8 is *highest*, D10 is *highest*, D6 and D8 are *highest*, D8 and D10 are *highest*, D6 and D10 are *highest* and finally all 3 dice are equal (and thus *highest*)". The probabilities over these cases would sum up to 1.