Karlsruhe Institute of Technology

Stefan Walzer, Stefan Hermann
ITI Prof. Sanders
winter term 2025/2026

Exercise Sheet 10
Approximation Algorithms

Probability and Computing

Exercise 1 – Jensen's Inequality

Let $D \subseteq \mathbb{R}$ be a connected domain and $f : D \to \mathbb{R}$ be a function. The function f is called convex if it is "curved to the left" and concave if it is "curved to the right".[1] A function is convex if and only if its negation is concave. For a formal definition see: Wikipedia

(a) Decide (without proof) for the following functions whether they are convex on their respective domains, concave, both, or neither.

$$f_1(x) = x, \qquad f_2(x) = x^2, \qquad f_3(x) = x^3, \qquad f_4(x) = \log(x), \qquad f_5(x) = \log^2(x).$$

(b) Let f be a convex function with domain D. Argue geometrically that for every $x_0 \in D$ there exists a linear function g such that:

(i) $f(x) \geq g(x)$ for all $x \in D$

(ii) $f(x_0) = g(x_0)$.

(c) Conclude that for every convex function f and for every random variable X with values in the domain D of f the following holds:

$$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X]).$$

Hint: Consider $x_0 = \mathbb{E}[X]$ and the corresponding g from the previous subproblem.

(d) Show that analogously, for every concave function f with domain D and for every random variable X with values in D the following holds:

$$\mathbb{E}[f(X)] \leq f(\mathbb{E}[X]).$$

The inequality from (c) as well as variants as in (d) are called Jensen's inequality.

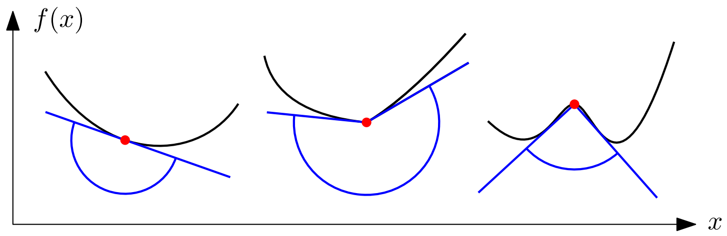
[1]The "curved to the left" in quotation marks allows, besides left curvatures (of a twice continuously differentiable function), also left kinks and linear behavior.

Solution 1

(a) The functions are all twice continuously differentiable. If the second derivative is everywhere non-negative, then the function is convex; if it is everywhere non-positive, then the function is concave.

- f_1 is convex and concave
- f_2 is convex
- f_3 is neither convex nor concave
- f_4 is concave
- f_5 is neither convex nor concave

(b) Because f is curved to the left, one can place a tangent g to the graph of f at the point $(x_0, f(x_0))$ such that f lies entirely above g.



That this is possible can be seen as follows (illustrated in the figure): One considers, below the point $(x_0, f(x_0))$ (red) on the graph of f (black), the angular region of those directions that never go beyond the graph of the function (blue). If this region is smaller than 180° (right in the figure), then one obtains a contradiction to the convexity of f. If this region is larger than 180° (center) or equal to 180° (left), then there exists at least one line through $(x_0, f(x_0))$ that avoids going beyond the graph of f.

(c) Let $g(x)$ be the function from (b) for $x_0 = \mathbb{E}[X]$. Then $f(x) \geq g(x)$ for all $x \in D$ and $f(\mathbb{E}[X]) = g(\mathbb{E}[X])$. Since g is a line, there exist $a, b \in \mathbb{R}$ such that $g(x) = ax + b$. It follows that

$$\mathbb{E}[f(X)] \geq \mathbb{E}[g(X)] = \mathbb{E}[aX + b] = a\mathbb{E}[X] + b = g(\mathbb{E}[X]) = f(\mathbb{E}[X]).$$

(d) Since $-f$ is convex, it follows directly from (c):

$$\mathbb{E}[f(X)] = -\mathbb{E}[-f(X)] \overset{(c)}{\leq} -(-f(\mathbb{E}[X])) = f(\mathbb{E}[X]).$$

Exercise 2 – Analysis of Lossy Counting

Reminder: Lossy Counting is a simple streaming algorithm that approximately counts the length m of a stream. It involves a parameter $p \in (0, 1]$. The algorithm itself as well as the way it is used are shown on the right. Prove:

(a) $\mathbb{E}[\text{result}] = m$

(b) $\Pr[|\text{result} - m| \le \varepsilon m] \ge 1 - 2\exp(-\varepsilon^2 pm/3)$.

(c) $\mathbb{E}[\text{space}] \le \log(1 + mp) + 1$.
Hint: By space we denote the maximum memory usage required for the state Z of LossyCounting. A number $i \in \mathbb{N}$ can be encoded with $\lceil \log_2(i+1) \rceil$ bits. Use Jensen's inequality from Exercise 1.

Algorithm init:
  $Z \leftarrow 0$
  return Z

Algorithm update(Z, a)**:**
  with probability p **do**
    $Z \leftarrow Z + 1$
  return Z

Algorithm result(Z)**:**
  return Z/p

Usage:
  $Z \leftarrow \text{init}()$
  for $i = 1$ **to** m **do**
    $Z \leftarrow \text{update}(Z, a_i)$
  return result(Z)

Solution 2

(a) Let $X_1, \dots, X_m \sim \text{Ber}(p)$ be independent random variables, where X_i indicates whether the i-th element of the stream leads to an increment of the counter Z. Then $X := \sum_{i=1}^{m} X_i$ is the value of Z after the last update. The estimate of the algorithm for m is thus result $= X/p$. Hence:

$$\mathbb{E}[\text{result}] = \mathbb{E}[X/p] = \tfrac{1}{p}\mathbb{E}\Big[\sum_{i=1}^{m} X_i\Big] = \tfrac{1}{p}\sum_{i=1}^{m}\mathbb{E}[X_i] = \tfrac{1}{p}\sum_{i=1}^{m} p = m.$$

(b) Using the stated Chernoff bound, we obtain

$$\begin{aligned}\Pr[|\text{result} - m| \ge \varepsilon m] &= \Pr[|X/p - m| \ge \varepsilon m] = \Pr[|X - mp| \ge \varepsilon mp] \\ &= \Pr[|X - \mathbb{E}[X]| \ge \varepsilon \mathbb{E}[X]] \le 2\exp(-\varepsilon^2 \mathbb{E}[X]/3) = 2\exp(-\varepsilon^2 mp/3).\end{aligned}$$

The claim follows by considering the complementary probability.

(c) Since Z grows monotonically, the memory requirement for Z is largest at the very end, namely $\lceil \log_2(1 + X) \rceil$. Since $f(x) = \log(1 + x)$ is concave on $[0, \infty)$, Jensen's inequality yields

$$\begin{aligned}\mathbb{E}[\text{space}] &= \mathbb{E}[\lceil \log_2(1 + X) \rceil] \le \mathbb{E}[\log_2(1 + X)] + 1 \\ &\overset{\text{Jensen}}{\le} \log_2(1 + \mathbb{E}[X]) + 1 = \log_2(1 + mp) + 1.\end{aligned}$$