# Problem 1: Optimization and probability

## 1a

$x_1, ..., x_n$ is a sequence of real numbers.
$w_1, ..., w_n$ are measuring the importance of numbers in a sequence
$L = \frac{1}{2}\sum_{i=1}^{n} w_i(\theta - x_i)^2$
$\frac{dL}{d\theta} = \sum_{i=1}^{n} w_i(\theta - x_i) = 0$
$\sum_{i=1}^{n} w_i\theta = \sum_{i=1}^{n} w_i x_i$
$\theta = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$
The problem arising from $w_i$'s being negative is that loss could potentially be unbounded from below.

## 1b

$f(x) = \sum max_{s_i \in \{-1,1\}} s_i x_i$
$g(x) = max_{s \in \{-1,1\}} \sum s x_i$
$g(x) = |\sum x_i|$
$f(x) = \sum |x_i|$
so $g(x) <= f(x)$

## 1c

Suppose you repeatedly roll a fair six-sided die until you roll a 1 (and then you stop). Every time you roll a 2, you lose a points, and every time you roll a 6, you win b points. You do not win or lose any points if you roll a 3, 4, or a 5. What is the expected number of points (as a function of a and b) you will have when you stop?
x - number of points
$E(x) = 1/6 * (E(x) - a) + 1/6 * (E(x) + b) + 3/6 * E(x)$
$\frac{E(x)}{6} = \frac{b-a}{6}$
$E(x) = b - a$

## 1d

Suppose the probability of a coin turning up heads is 0¡p¡1, and we flip it 7 times and get H,H,T,H,T,T,H. We know the probability (likelihood) of obtaining this sequence is $L(p) = p^4(1-p)^3$. What value of p maximizes L(p)? Prove/Show that this value of p maximizes L(p). What is an intuitive interpretation of this value of p?
$\log L(p) = 4\log p + 3\log(1-p)$, $\frac{d\log L(p)}{dp} = \frac{4}{p} - \frac{3}{1-p} = 0$, $\frac{4(1-p)}{1} = \frac{3p}{1}$, $p = \frac{4}{7}$
x is a small perturbation to our value of p
$(4/7 + x)^4(3/7 - x)^3 = -x^7 - x^6 + (3x^5)/7 + (29x^4)/49 - (16x^3)/343 - (288x^2)/2401 + 6912/823543$
First non-const term is negative quadratic in x, so it is a maximum indeed.

## 1e

Now for a little bit of practice manipulating conditional probabilities. Suppose that A and B are two events such that P($A|B$)=P($B|A$). We also know that and P($A \cup B$)=1 and P($A \cap B$) $> 0$. Prove that $P(A) > 1/2$.
P($A \cap B$) = P($A|B$)P(B)=P(A)
P($A \cap B$) = P($B|A$)P(A)=P(B)
P(B) = P(A)
1 = P($A \cup B$) = P(A) + P(B) - P($A \cap B$)=2 P(A) - P($A \cap B$) = 2P(A) - x, $x \in (0, 1]$
$P(A) = 1/2 + x/2 > 1/2$

## 1f

$$\nabla f_k = \frac{\partial f}{\partial w_k} = 2\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ik} - b_{jk}) + \frac{\lambda w_k}{\sqrt{\sum_{l=1}^{d} w_l^2}}$$

# Problem 2: Complexity

## 2a

Suppose we have an n×n grid, where we'd like to place 6 arbitrary axis-aligned rectangles (i.e., the sides of the rectangle are parallel to the axes). There are no constraints on the location or size of the rectangles. For example, it is possible for all four corners of a single rectangle to be the same point (resulting in a rectangle of size 0) or for all 6 rectangles to be on top of each other. How many possible ways are there to place 6 rectangles on the grid? In general, we only care about asymptotic complexity, so give your answer in the form of O(nc) or O(cn) for some integer c.
We have nxn grid. Every rectangle is fully specified by two x's and two y's. For one rectangle, we have (n*n)/2*(n*n)/2 = $n^4/4$. Twos in denominators are from double counting. So for six rectangles we would've $(n^4/4)^6 = O(n^{24})$

## 2b

Suppose we have an n×n grid. We start in the upper-left corner (position (1,1)), and we would like to reach the lower-right corner (position (n,n)) by taking single steps down or to the right. Suppose we are provided with a function c(i,j) that outputs the cost of touching position (i,j), and assume it takes constant time to compute for each position. Note that c(i,j) can be negative. Give an algorithm for computing the cost of the minimum-cost path from (1,1) to (n,n) in the most efficient way (with smallest big-O time complexity). What is the runtime (just give the big-O)?
Solution is provided in terms of dynamic programming.
cost(i,j) = min(cost(i-1,j),cost(i,j-1)), boundary conditions has to be taken into account. This algorithm works on $O(n^2)$ time