**Problem 1**

1. **Minimizing f(θ):** To find the value of θ that minimizes the function f(θ), which is defined as the sum of weighted squares, we can calculate θ using the following formula:

θ = (Σ(w\_i \* x\_i)) / (Σw\_i)

This θ represents the point where f(θ) is minimized.

1. **Showing it's a Minimum:** We can confirm that this θ is indeed a minimum by noting that the second derivative of f(θ) is always positive, indicating a minimum point.
2. **Issues with Negative *wi*​'s:** If some of the weights (w\_i) are negative, it can lead to unpredictable and unconventional behavior in the optimization problem. This can make it difficult to find a meaningful minimum. It's crucial to ensure that all weights are positive for a well-defined optimization problem.

**Problem 2**

We want to determine whether *f*(*x*)≤*g*(*x*) for all *x*.

First, let's understand *f*(*x*) and (*g*(*x*):

* *f*(*x*) finds the maximum of *sxi* for each *i* within the closed interval [−1,1] and multiplies these maximum values together.
* *g*(*x*) first calculates the maximum value of *xi* for each *i* within the same interval and then multiplies these maximum values together.

Now, consider this: for each *i*, *sxi* is maximized over the same interval [−1,1] in both functions. Therefore, the maximum value of *sxi* over this interval is always greater than or equal to the maximum value of *xi* over the same interval.

So, for each *i*, we have *sxi*≥*xi* for all *s* in [−1,1]. When we multiply all these inequalities for *i*=1 to *d*, we get the following:

*f*(*x*)=max*s*∈[−1,1]​(*sxi*)⋅max*s*∈[−1,1]​(*sxi*)⋅...⋅max*s*∈[−1,1]​(*sxi*)≤max*s*∈[−1,1]​*xi*⋅max*s*∈[−1,1]​*xi*⋅...⋅max*s*∈[−1,1]​*xi*=*g*(*x*)

So, for all *x*, *f*(*x*) is indeed less than or equal to *g*(*x*).

**Problem 3**

Let's define *V* as the expected number of points you will have when you stop playing the game.

When you roll a 1 or a 2, you stop, and your expected points are 0.

When you roll a 3, you lose a point (*a*) and then start again, so your expected points become *V*−*a*.

When you roll a 4 or a 5, nothing happens, and your expected points remain *V*.

When you roll a 6, you win *b* points and then start again, so your expected points become *V*+*b*.

Now, we can set up the recurrence for *V*=1/6(0)+1/6(0)+1/6(V - a)+1/6(V)+1/6(V + b)

Simplifying this equation:

*V*=61​(2*V*+*b*−*a*)

Now, solving for *V*:

6*V*=2*V*+*b*−*a*

4*V*=*b*−*a*

4*V*=4*b*−*a*​

So, the expected number of points you will have when you stop playing the game is (*b*−*a)/4*​.

**Problem 4**

Suppose the probability of a coin turning up heads is *p* (where 0<10<*p*<1), and we flip it 6 times and get the sequence {T, H, H, H, T, H}. We know the probability (likelihood) of obtaining this sequence is given by 2*L*(*p*)=*p*4(1−*p*)2. 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*?

**Solution:**

To find the value of *p* that maximizes *L*(*p*)=*p*4(1−*p*)2, we'll take the derivative of log(*L*(*p*)) with respect to *p* and set it equal to zero. This will help us identify the maximum point.

1. Start with the likelihood function *L*(*p*)=*p*4(1−*p*)2
2. Take the natural logarithm (log) of *L*(*p*): log(*L*(*p*)=log(*p*4(1−*p*)2)=4log(*p*)+2log(1−*p*)
3. Find the derivative of log⁡(�(�))log(*L*(*p*)) with respect to �*p*: ���[4log⁡(�)+2log⁡(1−�)]=4�−21−�*dpd*​[4log(*p*)+2log(1−*p*)]=*p*4​−1−*p*2​
4. Set the derivative equal to zero and solve for �*p*: 4�−21−�=0*p*4​−1−*p*2​=0

Solving for �*p*: 4(1−�)−2�=04(1−*p*)−2*p*=0 6�=46*p*=4 �=46*p*=64​ *p*=32​

So, the value of *p* that maximizes *L*(*p*) is *p*=32​.

To show that this value is indeed a maximum, we can take the second derivative of log⁡(�(�))log(*L*(*p*)) and confirm that it is negative at *p*=32​. This would indicate a local maximum point.

**Intuitive Interpretation:**

The value *p*=32​ maximizes the likelihood of obtaining the given sequence of coin flips {T, H, H, H, T, H}. Intuitively, this means that, based on the observed sequence, the most likely probability of getting heads in a single coin flip is 2332​, or 2 out of 3 times. In other words, the observed sequence suggests that the coin is biased towards landing heads more often than tails, with a probability of 2332​ for heads.