13/18

Maxim Firsov

1A.

2/2

1. Scalar Multiplication Associativity: For any scalar a and vectors u and v, (a * b) * v = a * (b * v). This means that scalar multiplication is associative.
2. Scalar Multiplication Distributivity (over vector addition): For any scalars a and b and a vector v, (a + b) * v = a * v + b * v. This axiom states that scalar multiplication distributes over vector addition.
3. Scalar Multiplication Distributivity (over scalar addition): For any scalar a and vectors u and v, a * (u + v) = a * u + a * v. This axiom implies that scalar multiplication distributes over scalar addition. → additive $-v+v=0$
4. Scalar ~~Multiplication~~ Identity: For any vector v, 1 * v = v. This axiom states that multiplying a vector by the scalar 1 yields the original vector.

1/1

1B.

In a vector space, scalar multiplication involves multiplying a vector by a scalar. The scalars need to exhibit properties such as closure, associativity, commutativity, existence of multiplicative identity, existence of multiplicative inverses (except for the zero scalar), and distributivity over vector addition. These properties align with the defining properties of a field.

2/2

2A.

In the n + 1-dimensional firing rule, the neuron fires when $\Sigma_{i=0}^{n} w_i x_i > 0$. Here, the sum includes an additional term $w_0 x_0$, where $x_0 = 1$. A bias, which has a negative value ($-w_0$).
For the n + 1-dimensional firing rule, when the sum $\Sigma_{i=0}^{n} w_i x_i$ is greater than 0, it implies that: $w_0 x_0 + \Sigma_{i=1}^{n} w_i x_i > 0$.
Since $x_0 = 1$: $w_0 + \Sigma_{i=1}^{n} w_i x_i > 0$.
Equal to: $\Sigma_{i=1}^{n} w_i x_i > -w_0$.
With the condition for the n-dimensional firing rule ($\Sigma_{i=1}^{n} w_i x_i > T$), we observe that the inequalities are equivalent, with T being equal to $-w_0$.
Thus the firing rules describe the same condition for the neuron to fire. The condition under an additional dimension is the same, only with bias.

1/1

2B.

H(n dimensions) does not include the origin (0 vector) as part of the hyperplane, while H_(n+1 dimensions) includes the origin.

2C and 2D.

1. Multiplying all coefficients by 5:
   - Hyperplane H: If all coefficients are multiplied by 5: $\Sigma_{i=1}^{n} (5w_i)x_i = -(5w_0)$.

1/2

The hyperplane H in n dimensions changes, but the firing rule remains the same since the ratio between the coefficients and the threshold (T) remains unchanged.

   - Hyperplane H_: If all coefficients are multiplied by 5: $\Sigma_{i=0}^{n} (5w_i)x_i = 0$. The ~~hyperplane H_ change~~s, but the firing rule remains the same since the ratio between the coefficients and the threshold (0) remains unchanged.

• hyperplanes don't change

• both change for -5

1/2

2. Multiplying all coefficients by -5:
    - Hyperplane H: If all coefficients are multiplied by -5: $\Sigma_{i=1}^{n} (-5w_i)x_i = -(-5w_0) = 5w_0$. The hyperplane H in n dimensions changes, and the firing rule also changes since the threshold (T) becomes $-5w_0$ instead of $-w_0$.
    - Hyperplane H_: If all coefficients are multiplied by -5: $\Sigma_{i=0}^{n} (-5w_i)x_i = 0$. The hyperplane H_ in n + 1 dimensions changes, but the firing rule remains the same since the ratio between the coefficients and the threshold (0) remains unchanged.

2D and 2F. — you never answer either question

1/1

0/1

- Hyperplane H: $\Sigma_{i=1}^{n} (w_i/wn)x_i = -(w_0/wn)$.

The coefficient $w_{\square}$ is no longer present in the hyperplane equation. As a result, the number of free parameters needed to specify the hyperplane is reduced by 1, as $w_{\square}/wn = 1$. the firing rule becomes $x_0 = 1$ and $\Sigma_{i=1}^{n} (w_i/wn)x_i > (T/wn)$. The threshold (T) also gets divided by wn to maintain the ratio between the sum of products and the threshold changing the rule.

- Hyperplane H_: $(w_i/wn)x_i = 0$.

The coefficient $w_{\square}$ is no longer present in the hyperplane equation of H_ in this "reduced parameter representation." The firing rule becomes $\Sigma_{i=0}^{n} (w_i/wn)x_i > 0$, keeping it unchanged.

3A.

2/2

For n = 1 (a line):
$\Xi(1, 0) = 1$ (no hyperplanes)
$\Xi(1, 1) = 2$ (1 hyperplane divides the line into two regions)
$\Xi(1, 2) = 3$ (2 hyperplanes divide the line into three regions)
$\Xi(1, 3) = 4$ (3 hyperplanes divide the line into four regions)

From that:
$\Xi(1, m) = m + 1$

For n = 2 (a plane), the formula becomes:
$\Xi(2, m) = (m + 1) + \Xi(1, m) = (m + 1) + (m + 1) = 2(m + 1)$

For n = 3 (3-dimensional space), the formula becomes:
$\Xi(3, m) = (m + 1) + \Xi(2, m) = (m + 1) + 2(m + 1) = 3(m + 1)$

Assuming the pattern continues:
$\Xi(n, m) = (n-1)(m + 1) + 1$

Thus, the explicit formula for $\Xi(1, m)$ is $m + 1$.

1/1

3B.

The formula is right because each hyperplane added to n-dimensional space introduces m + 1 new regions. This occurs because each hyperplane intersects the existing hyperplanes, creating new divisions. The (n-1) term accounts for the total effect of hyperplanes in each dimension. Finally, the additional "+1" term represents the original region before any hyperplanes were added.

1/3

3C.

As m becomes significantly larger compared to n, the contribution of m_n to the overall sum becomes the most significant. The other terms, $m_0, m_1, \ldots, m_{n-1}$, may still have an impact, but their growth is slower compared to m_n. Therefore, for large values of m, these terms can be considered less significant in the overall computation.

• do a proof