+1

19/18

clean notebook

2/2

Problem 1a
What axioms are the remaining axioms of a vector space?

Solution: Let $a, b, c \in \mathbb{F}$, while u and v are arbitrary elements of a set V. Then to be a vector space, we'll additionally require of our set:

u or v, not a

1. Distributively (over both vector and field addition), e.g.,

$$a(u+v) = au + av \quad \text{and} \quad \cancel{a(b+c) = ab + ac}$$

u(b+c) = ub + uc

2. Multiplicative Identity element $1_v \in \mathbb{F}$ such that

$$1_v u = u$$

3. Vector Inverses. For each element $v \in V$, we require a negation $-v \in V$ such that

$$-v + v = 0 \quad \forall v \in V$$

4. A multiplication operation is defined between elements of our field and elements of the vector space, E.g.,

$$a(bv) = (ab)v$$

1/1

Problem 1b
Are scalars in a vector space from groups, rings, or fields?

Solution: We require our scalars come from a field. As summarized from wikipedia: "a field has two operations, called addition and multiplication; it is an abelian group under addition with 0 as the additive identity; the nonzero elements are an abelian group under multiplication with 1 as the multiplicative identity; and multiplication distributes over addition."

Note, being abelian is what people usually call "commutativity". As such, fields satisfy the exact requirements outlined in problem 1a. This reasoning is a bit circular because vector spaces *were constructed* to be over fields. They are called modules when taken over rings.

2/2

Problem 2a
Describe why the hyperplanes H and $\underline{H}$ represent the same firing rule for the neuron?

Solution: Primarily, the first plane $\underline{H}$ is defined by the equation

$$\sum_{i=1}^{n} w_i x_i = -w_0 \iff w_1 x_1 + \ldots + w_n x_n = -w_0$$

while H is defined by

$$\sum_{i=0}^{n} w_i x_i = 0 \iff w_0 x_0 + w_1 x_1 + \ldots + w_n x_n = 0.$$

Since $x_0 = 1$ we can rewrite the above to give

$$w_1 x_1 + \ldots + w_n x_n = -w_0$$

which is precisely the equation that defined H. Thus this is simply an embedding of H in an additional dimension, except our plane is additionally shifted w_0 to the origin. As such, the firing rule which depended on an inner product of our weight and state vector being being greater than $-w_0$ is now equivalent to the firing rule that requires our inner product being greater than 0 after the shift.

1/1

Problem 2b
Assume $w_0 \neq 0$, what is the special property of H that $\underline{H}$ lacks?

Solution: H goes through the origin (has right hand side equal to zero)!

Problem 2c
Multiply the weights by ± 5, what happens?

2/2

Solution: Neither hyperplane changes after the weights are multiplied by 5, which can be seen below:

$$5w_0x_0 + 5w_1x_1 + \ldots + 5w_nx_n = 0 \implies w_0x_0 + w_1x_1 + \ldots + w_nx_n = 0$$

after division by 5, which is still the same equation that defined H. For $\underline{H}$ we also see no change:

$$5w_1x_1 + \ldots + 5w_nx_n = -5w_0 \implies w_1x_1 + \ldots + w_nx_n = -w_0$$

again after division by 5. This is the same equation that defined $\underline{H}$. The same argument holds for -5 as well, and in fact any constant c.

Problem 2d
Does either firing rule change?

2/2

Solution: When multiplied by 5 the firing rules are unchanged, assuming the threshold is scaled appropriately. To make this more clear, anything that would trigger a fire for H has inner product strictly greater than zero. Anything strictly greater than zero multiplied by 5 is still strictly greater than zero.

However, when multiplied by -5, the firing rules flip. For example

$$\sum_{i=0}^{n} w_ix_i > 0 \implies \sum_{i=0}^{n} -5w_ix_i < 0$$

with a similar situation holding for the firing rule of $\underline{H}$.

1/1

Problem 2e
Suppose $w_0, \ldots, w_n$ are all divided by w_n. Does either hyperplane change under this reduced parameter representation?

Solution: No, neither hyperplane changes. The exact argument in 2c holds in general, that is one can replace 5 with any constant (such as $\frac{1}{w_n}$).

1/1

Problem 2f
Do either of the firing rules change?

Solution: The firing rule changing depends on the sign of w_n. If w_n is positive, the rule stays the same. If w_n is negative, the rule flips.

2/2

Problem 3a
Give an explicit formula for $\Xi(1, m)$.

Solution: Consider the formula

$$\Xi(1, m) = m + 1$$

1/1

Problem 3b
Geometrically, why is this formula right?

Solution: Consider "1-space", a line. Placing a hyperplane here is equivalent to making a cut on the line, which is known to result in 1 more segment than the number of cuts made. Hence we arrive at $m + 1$ for m the number of cuts (or hyperplanes).

there is a less convoluted way to show this (slightly less rigor)

3/3

Problem 3c

Show

$$\sum_{i=0}^{n} \binom{m}{i} = O(m^n)$$

for fixed n and variable m.

Primarily recall the definition of big O. I will argue

$$\lim_{m\to\infty} \sup \frac{|f(m)|}{g(m)} < \infty \tag{1}$$

where $f(m) = \sum_{i=0}^{n} \binom{m}{n}$ and $g(m) = m^n$. Since n is fixed, our function f will continue to monotonically increase as m increases, eventually being dominated by the greatest term in the summation. Observe since $m >> n$, it is safe to assume $n \leq \frac{m}{2}$. Then this means our summation defined in f is dominated by the final term in the summation (due to the symmetry of binomial coefficients), which we may call

$$\binom{m}{n} = \frac{m!}{n!(m-n)!}$$

for large enough m. Then equation (1) can be rewritten as

$$\begin{aligned}\lim_{m\to\infty} \sup \frac{|f(m)|}{g(m)} &= \lim_{m\to\infty} \frac{m!}{n!(m-n)!}\left(\frac{1}{m^n}\right) \\ &= \lim_{m\to\infty} \frac{(m-1)(m-2)\dots(m-n+1)}{n!}\left(\frac{1}{m^n}\right) < \infty\end{aligned} \tag{2}$$

as desired. Thus we conclude $f = O(g)$ as desired.