Homework 2

2.160.)

a.

$$\begin{bmatrix} x_1 \\ 0.5(x_1+x_2) \\ x_2 \\ 0.5(x_2+x_3) \\ x_3 \\ 0.5(x_3+x_4) \\ \vdots \\ x_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\Rightarrow A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & ... & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & 0 & 0 & ... & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 & ... & 0 \\ 0 & 0 & 1 & 0 & 0 & ... & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & ... & 1 \end{bmatrix}$$

b.

$$\begin{bmatrix} x_2 \\ x_4 \\ x_6 \\ \vdots \\ x_n \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\Rightarrow A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & ... & 1 \end{bmatrix}$$

c.

$$\begin{bmatrix} 0.5(x_1+x_2) \\ 0.5(x_3+x_4) \\ 0.5(x_5+x_6) \\ \vdots \\ 0.5(x_{n-1}+x_n) \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\Rightarrow A = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & ... & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & ... & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & ... & 0.5 & 0.5 \end{bmatrix}$$

3.250.)

a.

We can represent this problem with the following matrix multiplication:

$$y = Ap$$

$$\begin{bmatrix} L_{cone} \\ M_{cone} \\ S_{cone} \end{bmatrix} = \begin{bmatrix} l_1 & l_2 & ... & l_{20} \\ m_1 & m_2 & ... & m_{20} \\ s_1 & s_2 & ... & s_{20} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{20} \end{bmatrix}$$

Where $p = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{20} \end{bmatrix}$ is the given light spectra (the input), $A = \begin{bmatrix} l_1 & l_2 & ... & l_{20} \\ m_1 & m_2 & ... & m_{20} \\ s_1 & s_2 & ... & s_{20} \end{bmatrix}$ describes each cone's spectral response to every band, and $y = \begin{bmatrix} L_{cone} \\ M_{cone} \\ S_{cone} \end{bmatrix}$ is the net spectral responses of all the cones (the output). Since color perception only depends on the spectral responses of the cones, y, that means two light spectra, p and $\tilde{p}$, are visually indistinguishable when they have the same output $y = \tilde{y}$. From here, we get

$$y = \tilde{y} \Leftrightarrow Ap = A\tilde{p} \Leftrightarrow A(p - \tilde{p}) = 0$$

In other words, two light spectra, p and $\tilde{p}$, are visually indistinguishable when $(p - \tilde{p}) \in \text{null}(A)$.
*It's worth noting that there might be cases where $y \neq \tilde{y}$, but the two light spectra are visually indistinguishable because we don't know the complex function transforming the spectral responses of the cones into the perceived color. However, we described all the solutions we know about without knowing said complex function.

b.

This question can be rephrased by using the problem setup described in part a with the equation below:

$$Ap_{\text{test}} = Ap_{\text{match}} = A(a_1u + a_2v + a_3w)$$

The question asks for any arbitrary light spectra p is there guaranteed to be weights a_1, a_2, a_3 such that the equation holds true. To answer this, we first solve for a_1, a_2, a_3

$$A(a_1u + a_2v + a_3w) = A \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} Au & Av & Aw \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = B \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Where $B = \begin{bmatrix} Au & Av & Aw \end{bmatrix} \in \mathbb{R}^{3\times3}$. From here, we can solve for a_1, a_2, a_3 given any p if and only if B is invertible. If so, our solution would be

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = B^{-1}Ap_{\text{test}}$$

If B isn't invertible then $B \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ spans two dimensions at most, which isn't enough to cover the three dimension range of Ap. Whether or not we can always make p_{match} visually indistinguishable from p_{test} is the same as asking is B an invertible matrix or in other words, is the set $\{Au, Av, Aw\}$ linearly independent? We can't answer this without knowing A, u, v, w, but from context I believe the set is linearly independent because u, v, w are the spectra of the primary lights and I imagine the cones in eyes don't perceive one primary color as a linear combination of the other two.

c.

This question was solved using Julia, the source code will be included in the submission. Assuming u is the spectra for red, v is the spectra for green, and w is the spectra for blue, weights that achieve the match are

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0.4226 \\ 0.0987 \\ 0.5286 \end{bmatrix}$$

d.

Beth is right. Let I denote the light spectrum for a tungsten bulb, I' denote the light spectrum for sunlight, r denote the reflectance for one object and $\tilde{r}$ denote the reflectance for a different object. Then it's possible to have

$$A\begin{bmatrix} I_1 r_1 \\ I_2 r_2 \\ \vdots \\ I_{20} r_{20} \end{bmatrix} = A\begin{bmatrix} I_1 \tilde{r}_1 \\ I_2 \tilde{r}_2 \\ \vdots \\ I_{20} \tilde{r}_{20} \end{bmatrix} \Rightarrow A(\begin{bmatrix} I_1 r_1 \\ I_2 r_2 \\ \vdots \\ I_{20} r_{20} \end{bmatrix} - \begin{bmatrix} I_1 \tilde{r}_1 \\ I_2 \tilde{r}_2 \\ \vdots \\ I_{20} \tilde{r}_{20} \end{bmatrix}) = 0 \Rightarrow (\begin{bmatrix} I_1 r_1 \\ I_2 r_2 \\ \vdots \\ I_{20} r_{20} \end{bmatrix} - \begin{bmatrix} I_1 \tilde{r}_1 \\ I_2 \tilde{r}_2 \\ \vdots \\ I_{20} \tilde{r}_{20} \end{bmatrix}) \in \text{null}(A)$$

$$A\begin{bmatrix} I'_1 r_1 \\ I'_2 r_2 \\ \vdots \\ I'_{20} r_{20} \end{bmatrix} \neq A\begin{bmatrix} I'_1 \tilde{r}_1 \\ I'_2 \tilde{r}_2 \\ \vdots \\ I'_{20} \tilde{r}_{20} \end{bmatrix} \Rightarrow (\begin{bmatrix} I'_1 r_1 \\ I'_2 r_2 \\ \vdots \\ I'_{20} r_{20} \end{bmatrix} - \begin{bmatrix} I'_1 \tilde{r}_1 \\ I'_2 \tilde{r}_2 \\ \vdots \\ I'_{20} \tilde{r}_{20} \end{bmatrix}) \notin \text{null}(A)$$

Since $(\begin{bmatrix} I_1 r_1 \\ I_2 r_2 \\ \vdots \\ I_{20} r_{20} \end{bmatrix} - \begin{bmatrix} I_1 \tilde{r}_1 \\ I_2 \tilde{r}_2 \\ \vdots \\ I_{20} \tilde{r}_{20} \end{bmatrix})$ is in the null space of A, that means the two objects are visually identical when illuminated by a tungsten bulb and since $(\begin{bmatrix} I'_1 r_1 \\ I'_2 r_2 \\ \vdots \\ I'_{20} r_{20} \end{bmatrix} - \begin{bmatrix} I'_1 \tilde{r}_1 \\ I'_2 \tilde{r}_2 \\ \vdots \\ I'_{20} \tilde{r}_{20} \end{bmatrix})$ is not in the null space of A, that means the two objects are visually different when illuminated by sunlight. We found an example pair of objects that meets these criteria using the given tungsten and sunlight vectors in Julia

$$r = [0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5]^\intercal$$

$$\tilde{r} = [0.490, 0.492, 0.487, 0.516, 0.498, 0.499, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50, 0.50]^\intercal$$

3.260.)

We'll show that the set of points in $\mathbb{R}^n$ that are closer to a than b is a halfspace, i.e.

$$\{x \mid \|x-a\| \leq \|x-b\|\} = \{x \mid c^\intercal x \leq d\}$$

where c and d are defined as

$$c = b - a$$

$$d = \frac{\|b\|^2 - \|a\|^2}{2} = (b-a)^\intercal \frac{(b+a)}{2}$$

We'll do this with a chain of equivalences. Start with

$$\|x-a\| \leq \|x-b\|$$

Another way to rewrite this is

$$\sqrt{\langle x-a, x-a\rangle} \leq \sqrt{\langle x-b, x-b\rangle}$$

Since norms are always greater than or equal to 0 (don't have to worry about negatives for square root domain) and the square root function is monotonic and increasing, we know that

$$\sqrt{\langle x-a, x-a\rangle} \leq \sqrt{\langle x-b, x-b\rangle} \Leftrightarrow \langle x-a, x-a\rangle \leq \langle x-b, x-b\rangle$$

We can then rewrite this as

$$\langle x, x\rangle - \langle x, a\rangle - \langle a, x\rangle + \langle a, a\rangle \leq \langle x, x\rangle - \langle x, b\rangle - \langle b, x\rangle + \langle b, b\rangle$$

This then simplifies to

$$\|a\|^2 - 2\langle a, x\rangle \leq \|b\|^2 - 2\langle b, x\rangle$$

And further down to

$$2\langle b-a, x\rangle \leq \|b\|^2 - \|a\|^2$$

And end by rewriting this as

$$(b-a)^\intercal x \leq \frac{\|b\|^2 - \|a\|^2}{2}$$

We've thus proven the statement above. The directional (normal) vector for the halfspace is $c = b - a$ and a point on the halfspace is the midpoint between a and b, $\frac{b+a}{2}$, because $d = (b-a)^\intercal \frac{(b+a)}{2}$. This is shown in the drawing below:

3.300.)

a.

To prove $\mathcal{V}^\perp$ is a subspace of $\mathbb{R}^n$, we need to show that $\mathcal{V}^\perp$ is closed under scalar multiplication and vector addition. Fix $x \in \mathcal{V}^\perp$ and $c \in \mathbb{R}$. Fix $y \in \mathcal{V}$. Then

$$\langle x, y\rangle = 0 \Rightarrow \langle cx, y\rangle = c\langle x, y\rangle = 0$$

As y was arbitrary, that means $\langle cx, y\rangle = 0$ for all $y \in \mathcal{V}$, so $cx \in \mathcal{V}^\perp$. As x and c were arbitrary, that means $\mathcal{V}^\perp$ is closed under scalar multiplication. Now fix $u, v \in \mathcal{V}^\perp$ and $y \in \mathcal{V}$. Then

$$\langle u, y\rangle = 0 \wedge \langle v, y\rangle = 0 \Rightarrow \langle u+v, y\rangle = \langle u, y\rangle + \langle v, y\rangle = 0$$

As y was arbitrary, that means $\langle u+v, y\rangle = 0$ for all $y \in \mathcal{V}$, so $u + v \in \mathcal{V}^\perp$. As u and v were arbitrary, that means $\mathcal{V}^\perp$ is closed under vector addition. Hence, $\mathcal{V}^\perp$ is a subspace of $\mathbb{R}^n$.

b.

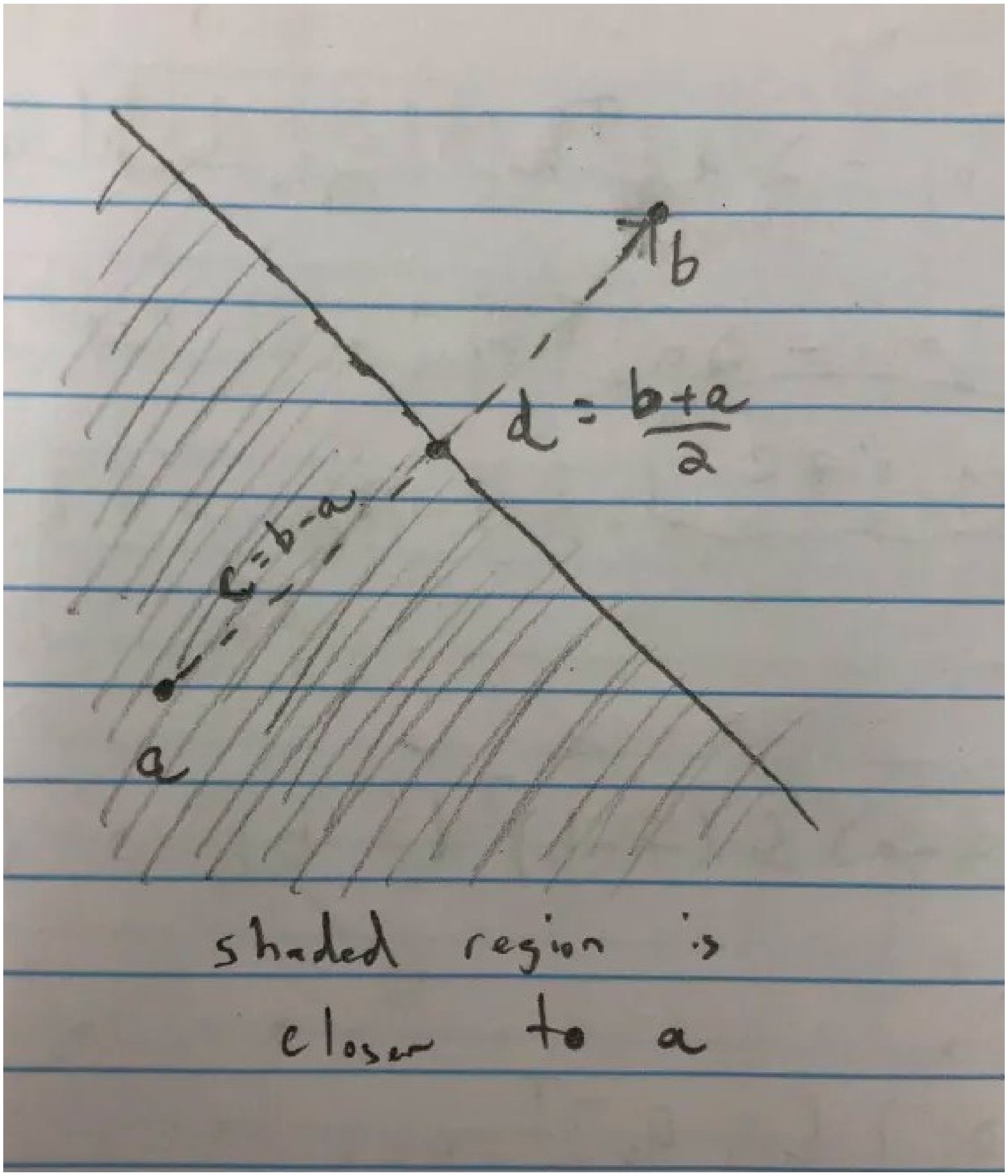
As $\mathcal{V}$ is the span of the vectors $v_1, v_2, ..., v_r$ and these same vectors form the columns of the matrix V, that means $\mathcal{V}$ is the range of V. For defining $\mathcal{V}^\perp$, notice that for each column v_i of V and for every $x \in \mathcal{V}^\perp$ that $v_i^\intercal x = 0$. This means that x is in the null space of $V^\intercal$, so $\mathcal{V}^\perp$ is a subspace of the null space of $V^\intercal$. Notice that if y is in the null space of $V^\intercal$ then y must be orthogonal to all of $v_1, v_2, ..., v_r$, so $y \in \mathcal{V}^\perp$. Hence, $\mathcal{V}^\perp$ is the null space of $V^\intercal$.

c.

Let $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ be an orthogonal matrix in $\mathbb{R}^n$ where Q_1 is an orthonormal basis for $\mathcal{V}$ and Q_2 is an orthonormal basis for $\mathcal{V}^\perp$. Then for any $x \in \mathbb{R}^n$, we have

$$x = Q_1 Q_1^\intercal x + Q_2 Q_2^\intercal x$$

Let $v = Q_1 Q_1^\intercal x$ and $v^\perp = Q_2 Q_2^\intercal x$. This means $v \in \mathcal{V}$ because $v = Q_1 y$ (where $y = Q_1^\intercal x$), so v is in the range of Q_1 which is just $\mathcal{V}$ and you can make a similar argument for $v^\perp \in \mathcal{V}^\perp$. Then we have $x = v + v^\perp$ such that $v \in \mathcal{V}$ and $v^\perp \in \mathcal{V}^\perp$. The reason that v and $v^\perp$ are unique is because v must be the projection of x onto $\mathcal{V}$ and $v^\perp$ must be the projection of x onto $\mathcal{V}^\perp$ and since projection is a function, the projection of a vector onto a space must be unique. Thus, every $x \in \mathbb{R}^n$ can be uniquely expressed as $x = v + v^\perp$ where $v \in \mathcal{V}$ and $v^\perp \in \mathcal{V}^\perp$.



b
d = (b+a)/2
c = b-a
a
shaded region is
closer to a

d.

Recall from part b that we have $\mathcal{V} = \text{range}(V)$ and $\mathcal{V}^{\perp} = \text{null}(V^{\intercal})$. Also note that since the rank of V is the same as the rank of $V^{\intercal}$, that means $\dim(\text{range}(V)) = \dim(\text{range}(V^{\intercal}))$.

$$n = \dim(\text{range}(V^{\intercal})) + \dim(\text{null}(V^{\intercal})) = \dim(\text{range}(V)) + \dim(\text{null}(V^{\intercal})) = \dim \mathcal{V} + \dim \mathcal{V}^{\perp}$$

e.

Assume that $\mathcal{V} \subseteq \mathcal{U}$. We want to show that $\mathcal{U}^{\perp} \subseteq \mathcal{V}^{\perp}$. To do this, we just have to show $x \in \mathcal{U}^{\perp} \Rightarrow x \in \mathcal{V}^{\perp}$. Fix $x \in \mathcal{U}^{\perp}$. By definition, this means that for all $y \in \mathcal{U}$, $\langle x, y \rangle = 0$. Since $\mathcal{V} \subseteq \mathcal{U}$, the previous statement implies $\langle x, y \rangle = 0$ for all $y \in \mathcal{V}$ as well and by definition that means that $x \in \mathcal{V}^{\perp}$. Hence, $\mathcal{V} \subseteq \mathcal{U} \Rightarrow \mathcal{U}^{\perp} \subseteq \mathcal{V}^{\perp}$.

3.450.)

We want to find some conditions on the constellation, $a_1, a_2, ..., a_n$, that would guarantee the minimum distance and maximum correlation decoding functions have the same output for any input a_{recd}. In other words, we want to find some conditions on $a_1, a_2, ..., a_n$ such that for any a_{recd},

$$\min_{i \in \{1,2,\ldots,n\}} \|a_{\text{recd}} - a_i\| = \max_{i \in \{1,2,\ldots,n\}} a_{\text{recd}}^{\intercal} a_i$$

To find these conditions, we start by breaking down the left side.

$$\min_{i \in \{1,2,\ldots,n\}} \|a_{\text{recd}} - a_i\| = \min_{i \in \{1,2,\ldots,n\}} \sqrt{(a_{\text{recd}} - a_i)^{\intercal}(a_{\text{recd}} - a_i)} = \min_{i \in \{1,2,\ldots,n\}} (a_{\text{recd}} - a_i)^{\intercal}(a_{\text{recd}} - a_i)$$

*we got rid of the square root by using the logic as in question 3.260

$$= \min_{i \in \{1,2,\ldots,n\}} a_{\text{recd}}^{\intercal} a_{\text{recd}} + a_i^{\intercal} a_i - 2a_{\text{recd}}^{\intercal} a_i = \min_{i \in \{1,2,\ldots,n\}} a_i^{\intercal} a_i - 2a_{\text{recd}}^{\intercal} a_i$$

*Note that $a_{\text{recd}}^{\intercal} a_{\text{recd}}$ is the same for any i, so we can remove it from the expression. To restate the question, we want to find some conditions on $a_1, a_2, ..., a_n$ such that for any a_{recd},

$$\min_{i \in \{1,2,\ldots,n\}} a_i^{\intercal} a_i - 2a_{\text{recd}}^{\intercal} a_i = \max_{i \in \{1,2,\ldots,n\}} a_{\text{recd}}^{\intercal} a_i$$

This will be true if $a_1^{\intercal} a_1 = a_2^{\intercal} a_2 = ... = a_n^{\intercal} a_n$ or equivalently if $\|a_1\| = \|a_2\| = ... = \|a_n\|$. For a specific counterexample, consider a simple case in $\mathbb{R}$ where $a_1 = 10$, $a_2 = -1$ and $a_{\text{recd}} = 1$. Then the minimum distance decoder would output a_2 as -1 is closer to 1 than 10 is, but the maximum correlation decoder would output a_1 as 10 and 1 are going in the same direction (they're both positive) whereas -1 and 1 are going in opposite directions (one positive and one negative).