Review Worksheet 2

1. Cruise Control

Suppose that we're working with a more advanced version of the robot car we built in the lab. Its state at timestep k is n dimensional, captured in $\vec{x}[k] \in \mathbb{R}^n$. The control at each timestep $\vec{u}[k] \in \mathbb{R}^m$. The system evolves according to the discrete-time equation

$$\vec{x}[k+1] = A\vec{x}[k] + B\vec{u}[k]. \tag{1}$$

We know the values of the $n \times n$ matrix A and the $n \times m$ matrix B (say for example estimated through system identification). For all parts, the initial condition is $\vec{x}[0] = \vec{0}$.

(a) We want to transform our system to a nicer set of coordinates in the S basis. S is an $n \times n$ invertible matrix. Let us write the transformed state as $\vec{z}[k] = S^{-1}\vec{x}[k]$ for all k. **Show that eq. (1) can be written in the form**

$$\vec{z}[k+1] = \widetilde{A}\vec{z}[k] + \widetilde{B}\vec{u}[k]. \tag{2}$$

with $\widetilde{A} = S^{-1}AS$ and $\widetilde{B} = S^{-1}B$. Show your work.

Solution:

$$\vec{x}[k+1] = A\vec{x}[k] + B\vec{u}[k] \tag{3}$$
$$S\vec{z}[k+1] = AS\vec{z}[k] + B\vec{u}[k] \tag{4}$$
$$\vec{z}[k+1] = \underbrace{S^{-1}AS}_{\widetilde{A}}\vec{z}[k] + \underbrace{S^{-1}B}_{\widetilde{B}}\vec{u}[k] \tag{5}$$

(b) **Prove that the system in eq. (2) is controllable if and only if the system in eq. (1) is controllable. Show your work.**

(HINT: Connect the controllability matrix of the system in eq. (2) to the controllability matrix of the system in eq. (1).)

Solution: We have

$$C_{\vec{z}} = \begin{bmatrix} \widetilde{B} & \widetilde{A}\widetilde{B} & \cdots & \widetilde{A}^{n-1}\widetilde{B} \end{bmatrix} \tag{6}$$

Note that

$$A^t = \left(S\widetilde{A}S^{-1}\right)^t = S\widetilde{A}S^{-1}S\widetilde{A}S^{-1}\left(S\widetilde{A}S^{-1}\right)^{t-2} = S\widetilde{A}^2S^{-1}\left(S\widetilde{A}S^{-1}\right)^{t-2} = \cdots \tag{7}$$
$$= S\widetilde{A}^tS^{-1}. \tag{8}$$

So now we can write the controllability matrix for eq. (1):

$$C_{\vec{x}} = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \tag{9}$$
$$= \begin{bmatrix} B & S\widetilde{A}S^{-1}B & \cdots & S\widetilde{A}^{n-1}S^{-1}B \end{bmatrix} \tag{10}$$
$$= \begin{bmatrix} SS^{-1}B & S\widetilde{A}S^{-1}B & \cdots & S\widetilde{A}^{n-1}S^{-1}B \end{bmatrix} \tag{11}$$
$$= S\begin{bmatrix} S^{-1}B & \widetilde{A}S^{-1}B & \cdots & \widetilde{A}^{n-1}S^{-1}B \end{bmatrix} \tag{12}$$
$$= S\begin{bmatrix} \widetilde{B} & \widetilde{A}\widetilde{B} & \cdots & \widetilde{A}^{n-1}\widetilde{B} \end{bmatrix} \tag{13}$$
$$= SC_{\vec{z}}. \tag{14}$$

Since S is a basis matrix, $\text{rank}(C_{\vec{x}}) = \text{rank}(C_{\vec{z}})$, and so $\text{rank}(C_{\vec{x}}) = n$ if and only if $\text{rank}(C_{\vec{z}}) = n$.

(c) Suppose (just for this problem subpart) that the system in eq. (1) is controllable, and define its controllability matrix as $C \in \mathbb{R}^{n \times mn}$. We want to reach a goal state $\vec{g} \in \mathbb{R}^n$ in exactly n timesteps; that is, we want $\vec{x}[n] = \vec{g}$. Recall $\vec{x}[0] = \vec{0}$. We define the sequence of minimum energy controls as $\vec{u}^\star = \begin{bmatrix} \vec{u}^\star[n-1] \\ \vdots \\ \vec{u}^\star[0] \end{bmatrix}$ where

$$\vec{u}^\star = \underset{\vec{u}}{\operatorname{argmin}} \|\vec{u}\|^2 \tag{15}$$

$$\text{s.t. } C\vec{u} = \vec{g}. \tag{16}$$

Prove that $\vec{u}^*$ is orthogonal to the nullspace of C.

(HINT: Consider a solution of $C\vec{u} = \vec{g}$ as $\vec{u}_{\text{sol}} = \vec{u}_{\text{null}} + \vec{u}_{\text{other}}$, where $\vec{u}_{\text{null}}$ is the component of $\vec{u}_{\text{sol}}$ in the nullspace of C, (i.e. $\vec{u}_{\text{null}}$ the projection of $\vec{u}_{\text{sol}}$ onto the nullspace of C). In this case, we would have that $\vec{u}_{\text{null}} \perp \vec{u}_{\text{other}}$)

Solution: Let us suppose that we found some $\vec{u}_{\text{sol}}$ that satisfies (16). We decompose $\vec{u}_{\text{sol}}$ into a component $\vec{u}_{\text{null}} \in \text{Null}(C)$ and a component $\vec{u}_{\text{other}} \perp \text{Null}(C)$ such that

$$\vec{u}_{\text{sol}} = \vec{u}_{\text{null}} + \vec{u}_{\text{other}}. \tag{17}$$

Plugging this into (16) gives us

$$\begin{aligned} \vec{g} &= C\vec{u}_{\text{sol}} & (18)\\ &= C(\vec{u}_{\text{null}} + \vec{u}_{\text{other}}) & (19)\\ &= C\vec{u}_{\text{null}} + C\vec{u}_{\text{other}} & (20)\\ &= C\vec{u}_{\text{other}}. & (21) \end{aligned}$$

The implication above tells us that $\vec{u}_{\text{other}}$ also satisfies (16).

Now we can consider the norm squared of $\vec{u}_{\text{sol}}$

$$\begin{aligned} \|\vec{u}_{\text{sol}}\|^2 &= \vec{u}_{\text{sol}}^\top \vec{u}_{\text{sol}} & (22)\\ &= (\vec{u}_{\text{null}} + \vec{u}_{\text{other}})^\top (\vec{u}_{\text{null}} + \vec{u}_{\text{other}}) & (23)\\ &= \vec{u}_{\text{null}}^\top \vec{u}_{\text{null}} + \vec{u}_{\text{null}}^\top \vec{u}_{\text{other}} + \vec{u}_{\text{null}}^\top \vec{u}_{\text{other}} + \vec{u}_{\text{other}}^\top \vec{u}_{\text{other}}. & (24) \end{aligned}$$

Recall that $\vec{u}_{\text{null}}$ is orthogonal to $\vec{u}_{\text{other}}$ by definition, meaning that $\vec{u}_{\text{null}}^\top \vec{u}_{\text{other}} = 0$. So, we can clean up the expression above as

$$\begin{aligned} \|\vec{u}_{\text{sol}}\|^2 &= \vec{u}_{\text{null}}^\top \vec{u}_{\text{null}} + \vec{u}_{\text{other}}^\top \vec{u}_{\text{other}} & (25)\\ \implies \|\vec{u}_{\text{sol}}\|^2 &= \|\vec{u}_{\text{null}}\|^2 + \|\vec{u}_{\text{other}}\|^2 & (26) \end{aligned}$$

which is essentially the Pythagorean Theorem. If we are trying to minimize $\|\vec{u}_{\text{sol}}\|^2$ then we should set $\vec{u}_{\text{null}} = \vec{0}$ since $\vec{u}_{\text{other}}$ already solves (16). Thus, our final optimal solution is $\vec{u^*} = \vec{u}_{\text{other}}$, which is completely orthogonal to the nullspace of C.

(d) Now let us work in the standard basis, with the system in eq. (1). Suppose $n = 3$ and $m = 1$ (so that $A \in \mathbb{R}^{3\times 3}$, $B \in \mathbb{R}^3$, $\vec{x}[k] \in \mathbb{R}^3$, and $u[k] \in \mathbb{R}$). The SVD of the controllability matrix C is given as

$$C = \begin{bmatrix} \vec{w}_1 & \vec{w}_2 & \vec{w}_3 \end{bmatrix} \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \\ \vec{v}_3^\top \end{bmatrix}, \tag{27}$$

with $\alpha > \beta > 0$.

Is the system controllable? Justify your answer.

If the system *is* controllable, find a sequence of inputs $\vec{u} = \begin{bmatrix} u[2] & u[1] & u[0] \end{bmatrix}^\top$, such that $\vec{x}[3] = \vec{g}$, for a specific $\vec{g} \in \mathbb{R}^3$. (Here $\vec{u}$ should be a function of $\vec{g}$).

If the system *is not* controllable, find a $\vec{g} \in \mathbb{R}^3$ that is unreachable by the system, i.e. find $\vec{g}$ such that there is *no* sequence of inputs $\vec{u}$ that makes $\vec{x}[3] = \vec{g}$.

All answers for this problem part should be in terms of $\vec{w}_i$, $\vec{v}_i$, α, and β.

(HINT: Remember how the SVD is connected to the column space and null space of the matrix and that $\vec{x}[0] = \vec{0}$.)

Solution: The system is not controllable; we have $\text{rank}(C) = 2 < 3 = n$, since C has one zero singular value. Since the initial condition $\vec{x}[0] = \vec{0}$, we know that

$$\vec{x}[3] = A^3\vec{x}[0] + C\vec{u} = C\vec{u}. \tag{28}$$

Thus, the vectors that are reachable are exactly $\text{Col}(C)$. Since from the properties of SVD we know that $\text{Col}(C) = \text{Span}(\vec{w}_1, \vec{w}_2)$ and $\vec{w}_3 \perp \text{Span}(\vec{w}_1, \vec{w}_2)$, we know that $\vec{w}_3$ is unreachable.

(e) We continue the setup of the previous part, repeated here. We work in the standard basis, with the system in eq. (1). The SVD of the controllability matrix C is given as in (27), with $\alpha > \beta > 0$.

Let $H \subseteq \mathbb{R}^3$ be the vector subspace of inputs $\vec{u} = \begin{bmatrix} u[2] & u[1] & u[0] \end{bmatrix}^\top$ which set $\vec{x}[3] = \vec{0}$. **Give a basis for H.** Justify your answer.

All answers for this problem part should be in terms of $\vec{w}_i$, $\vec{v}_i$, α, and β. Show your work.

(HINT: Remember that $\vec{x}[0] = \vec{0}$ and $\vec{x}[3] = C\vec{u}$.)

Solution: Again, we know that

$$\vec{x}[3] = C\vec{u}. \tag{29}$$

Then $\vec{x}[3] = \vec{0}$ if and only if $\vec{u} \in \text{Null}(C) = \text{Span}(\vec{v}_3)$ — since this is the part of V that corresponds to the 0 singular value. Thus $\vec{v}_3$ will be the basis for H.

2. Stability

Consider the complex plane below, which is broken into non-overlapping regions A through H. The circle drawn on the figure is the unit circle $|\lambda| = 1$.

(a) Consider the continuous-time system $\frac{d}{dt}x(t) = \lambda x(t) + v(t)$ and the discrete-time system $y(t + 1) = \lambda y(t) + w(t)$.

In which regions can the eigenvalue λ be for a *stable* system? Fill out the table below to indicate *stable* regions. Assume that the eigenvalue λ does not fall directly on the boundary between two regions.

	A	B	C	D	E	F	G	H
Continuous Time System $x(t)$	○	○	○	○	○	○	○	○
Discrete Time System $y(t)$	○	○	○	○	○	○	○	○

Solution: For the continuous time system to be stable, we need the real part of λ to be less than zero. Hence, C,D,G,H satisfy this condition.

On the other hand, for the discrete time system to be stable, we need the norm of λ to be less than one. Hence, A,B,C,D satisfy this condition.

(b) Consider the continuous time system

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \lambda x(t) + u(t) \tag{30}$$

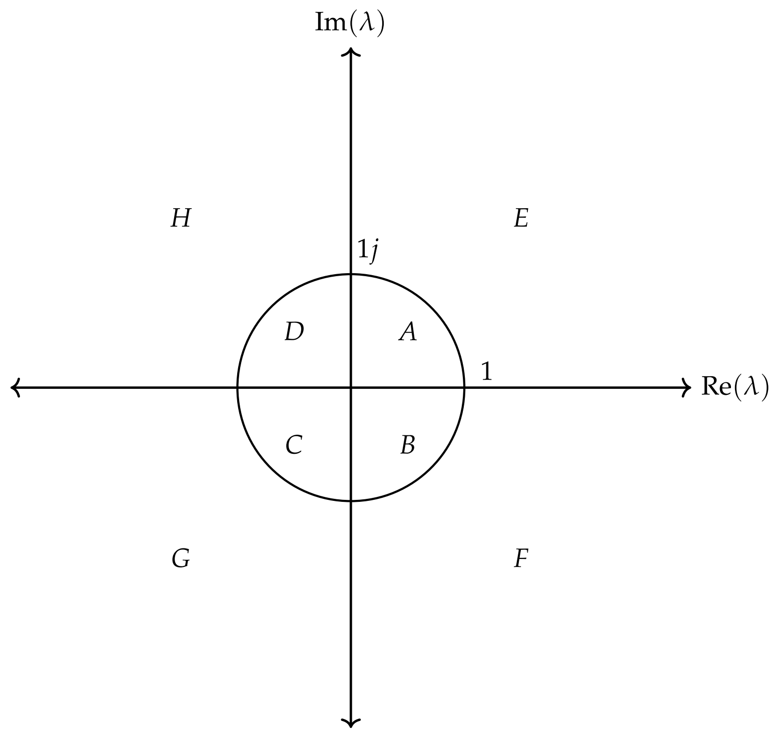

Figure 1: Complex plane divided into regions.

where λ is real and $\lambda < 0$. Assume that $x(0) = 0$ and that $|u(t)| < \epsilon$ for all $t \geq 0$. **Prove that the solution $x(t)$ will be bounded (i.e. $\exists k$ so that $|x(t)| \leq k\epsilon$ for all time $t \geq 0$).**
(HINT: Recall that the solution to such a first-order scalar differential equation is:

$$x(t) = x(0)e^{\lambda t} + \int_0^t u(\tau)e^{\lambda(t-\tau)}\, d\tau \tag{31}$$

You may use this fact without proof.)

Solution: Start by taking the abolute value of both sides.

$$|x(t)| = \left|x(0)e^{\lambda t} + \int_0^t u(\tau)e^{\lambda(t-\tau)}\, d\tau\right| \tag{32}$$

$$= \left|\int_0^t u(\tau)e^{\lambda(t-\tau)}\, d\tau\right| \tag{33}$$

$$\leq \int_0^t \left|u(\tau)e^{\lambda(t-\tau)}\right| d\tau \tag{34}$$

$$= \int_0^t |u(\tau)|\left|e^{\lambda(t-\tau)}\right| d\tau \tag{35}$$

$$< \int_0^t \epsilon\left|e^{\lambda(t-\tau)}\right| d\tau \tag{36}$$

$$= \epsilon \int_0^t \left|e^{\lambda(t-\tau)}\right| d\tau \tag{37}$$

$$= \epsilon \int_0^t e^{\lambda(t-\tau)}\, d\tau \tag{38}$$

$$= \epsilon \cdot \frac{e^{\lambda t} - 1}{\lambda} \tag{39}$$

$$= \epsilon \cdot \frac{1 - e^{\lambda t}}{-\lambda} \tag{40}$$

$$\leq \epsilon \cdot \frac{1}{-\lambda} \tag{41}$$

Hence the solution $x(t)$ will be bounded. In the above, we used the following integration:

$$\int_0^t e^{\lambda(t-\tau)}\, d\tau = e^{\lambda t} \cdot \int_0^t e^{-\lambda\tau}\, d\tau = e^{\lambda t}\left(-\frac{1}{\lambda}e^{-\lambda\tau}\right)_0^t = e^{\lambda t} \cdot \left(\frac{1}{\lambda} - \frac{1}{\lambda}e^{-\lambda t}\right) = \frac{e^{\lambda t} - 1}{\lambda} \tag{42}$$

3. I bet Cal will win this year

As huge fans of the Big Game, you and your friend want to bet on whether Cal or Stanford will win this year. You want to predict this year's result by analyzing historical records. Therefore, you decide to model this as a binary classification problem and do PCA for dimension reduction on the data you collected. The "+1" class represents victories of Cal and "−1" represents victories of Stanford.

After some research, you obtained a data matrix $A \in \mathbb{R}^{d \times n}$,

$$A = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix} \tag{43}$$

where each of the n columns $\vec{x}_i$ denotes a game and each of the d rows of A contains information of a possibly relevant factor of the games (weather, location, date, air quality, etc).

(a) Let the full SVD of $A = U\Sigma V^\top$, where A is given in eq. (43). You project your data along $\vec{u}_1$ and $\vec{u}_2$ (the first two principal components), and for comparison you also project your data along two randomly chosen directions $\vec{w}_1$ and $\vec{w}_2$ as well. You get the two pictures in Figure 2, but you forgot to label the axes. Of the two figures below, which one is the projection onto the principal components and which one is the projection onto the random directions? **Match axes (i), (ii), (iii), (iv) to $\vec{w}_1, \vec{w}_2, \vec{u}_1$, and $\vec{u}_2$, and justify your answer.**

Note that there may be multiple correct matchings; you only need to find and justify one of them.

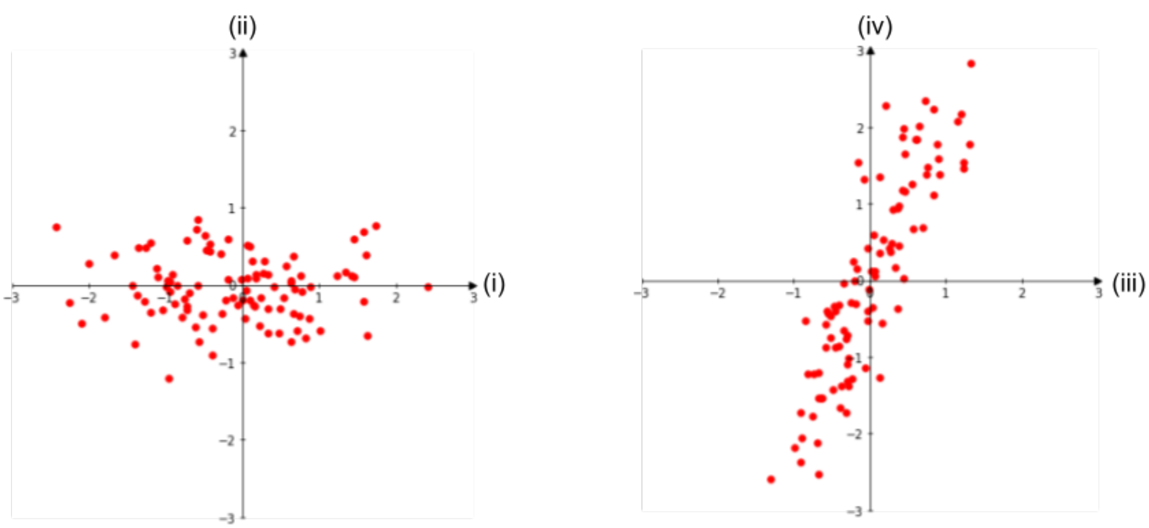

Figure 2: Projected datasets.

Solution: When we project data on the principal components, we require that one of the orthogonal axes maximizes the sum of squares of coordinates of the data points. We can see that the plot on the left is likely a projection onto the principal components, since the horizontal (i.e., x) axis maximizes the sum of the coordinates, which are the x-coordinates here. So we can deduce from this that (i) and (ii) must correspond in some order to the principal components and (iii) and (iv) must correspond to the random directions. Since axis (i) has a larger sum of squares than axis (ii) in the diagram on the left, it must be the case that axis (i) corresponds to $\vec{u}_1$, and the other corresponds to $\vec{u}_2$. The data is centered at (0, 0), so we can equivalently compare the "spread" along any given axis.

i. $\vec{u}_1$ – This is the axis with the maximal spread of the data and therefore must correspond to the largest singular value. the single axis, across both plots, across which there is maximal spread of the data.

ii. $\vec{u}_2$ – the corresponding axis to $\vec{u}_2$ and also an axis for which the spread of the data is axis-aligned.

iii. $\vec{w}_1$ or $\vec{w}_2$ – seemingly a random projection. We don't know which one is $\vec{w}_1$ and $\vec{w}_2$ since they are random unit vectors and as such are independent of the data, so we can't tell from the plot.

iv. $\vec{w}_2$ or $\vec{w}_1$ – seemingly a random projection. We don't know which one is $\vec{w}_1$ and $\vec{w}_2$ since they are random unit vectors and as such are independent of the data, so we can't tell from the plot.

(b) In order to reduce the dimension of the data, we would like to project the data onto the first k principal components of A, where k is less than the original data dimension d. **Show how to find the new vector $\vec{z}_i \in \mathbb{R}^k$ which is the k-dimesional, compressed version of $\vec{x}_i$.** You may use the SVD of A.

Solution: Write the matrix A in terms of its SVD, namely $A = U\Sigma V^\top$. In order to select the top k principal components, we can take the first k columns of U, denote U_k. To compress the data into k dimensions and find $\vec{z}_i$, we compute

$$\vec{z}_i = U_k^\top \vec{x}_i \tag{44}$$

(c) Given a new set of projection coefficients denoted $\vec{z}_{\text{new}} \in \mathbb{R}^k$, we can define a classifier that will predict $+1$ (i.e., that Cal wins) if $\vec{w}_\star^\top \vec{z}_{\text{new}} > 0$ and -1 (i.e., that Stanford wins) otherwise.

Assume $d = 6$, $k = 4$, and $\vec{w}_\star = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}^\top$. Let $A = U\Sigma V^\top$ for A defined in eq. (43), and you find that U is given by the identity matrix, i.e. $U = I_d$. Now suppose the data point for this year's big game $\vec{x}_{2021} = \begin{bmatrix} 3 & 6 & 4 & 1 & 9 & 6 \end{bmatrix}^\top$. **Would you bet on Cal or Stanford to win? Justify your answer.**

(HINT: Don't forget to project your data onto the principal components.)

Solution: First, we need to project the new data point onto the k-dimensional subspace just like what we did to the original data

$$\vec{z}_{2021} = U_k^\top \vec{x}_{2021} \tag{45}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^\top \begin{bmatrix} 3 \\ 6 \\ 4 \\ 1 \\ 9 \\ 6 \end{bmatrix} \tag{46}$$

$$= \begin{bmatrix} 3 \\ 6 \\ 4 \\ 1 \end{bmatrix} \tag{47}$$

Then, we compute the classifier's predicted value

$$p_{2021} = \vec{w}_\star^\top \vec{z}_{2021} \tag{48}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 4 \\ 1 \end{bmatrix} \quad (49)$$

$$= 6 > 0 \quad (50)$$

Therefore, the classifier predicts the label for this data point to be "+1", thus you should bet for Cal to win this year!

