Linear Systems: Assignment

Vectors

1. Is this set of vectors $\left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ independent? Explain your answer.

2. Consider a set of finite duration discrete-time real signals $X_N = \left\{ x[n] \,\middle|\, x[n] \in \mathbb{R}, \forall 0 \leq n \leq N-1 \right\}$. Does this set form a vector space? Explain your answer. Would X_N still be a vector spaces if the signals were binary signals? i.e. $x[n] \in \mathbb{B}$, where $\mathbb{B} = \{0, 1\}$ with the binary addition and multiplication operations defined as the following,

a	b	$a+b$	$a \times b$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Table 1: Addition and Multiplication operation for binary numbers.

3. Prove the following for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
   (a) **Triangle Inequality**:
   $$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$
   (b) **Backward Triangle Inequality**:
   $$\|\mathbf{x} - \mathbf{y}\| \geq |\|\mathbf{x}\| - \|\mathbf{y}\||$$
   (c) **Parallelogram Identity**:
   $$\frac{1}{2}\left(\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2\right) = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

4. Consider a set of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. When is $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} + \mathbf{y}\|$? What can you say about the geometry of the vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{x} - \mathbf{y}$ and $\mathbf{x} + \mathbf{y}$?

5. If $S_1, S_2 \subseteq V$ are subspaces of V, the is $S_1 \cap S_2$ a subspace? Demonstrate your answer.

6. Consider two sets of vectors $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\}$ and $W = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n, \mathbf{u}\}$. Prove that if $span(V) = span(W)$, then $\mathbf{u} \in span(V)$.

7. Prove that the sum of two subspaces $S_1, S_2 \subseteq V$ is a subspace.

8. Consider a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$. Express the following in-terms of inner product between a constant vector $\mathbf{u}$ and the given vector $\mathbf{v}$, and in each case specify the vector $\mathbf{u}$.
   (a) $\sum_{i=1}^{n} v_i$
   (b) $\frac{1}{n}\sum_{i=1}^{n} v_i$
   (c) $\sum_{i=1}^{n} v_i a^{(n-i)}$, where $a \in \mathbb{R}$
   (d) $\frac{1}{n-1}\sum_{i=1}^{n}\left(v_i - \frac{1}{n}\sum_{i=1}^{n} v_i\right)^2$
   (e) $\frac{1}{5}\sum_{i=3}^{5} v_i$
   (f) $\sum_{i=1}^{n-1}(v_{i+1} - v_i)$

9. Which of the following are linear functions of $\{x_1, x_2, \ldots, x_n\}$?
   (a) $\min_i \{x_i\}_{i=1}^{n}$
   (b) $\left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$
   (c) x_6

10. Consider a linear function $f : \mathbb{R}^n \to \mathbb{R}$. Prove that every linear function of this form can be represented in the following form.
$$y = f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} = \sum_{i=1}^{n} w_i x_i, \quad \mathbf{x}, \mathbf{w} \in \mathbb{R}^n$$

11. An *affine* function f is defined as the sum of a linear function and a constant. It can in general be represented in the form,
$$y = f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + \beta, \quad \mathbf{x}, \mathbf{w} \in \mathbb{R}^n, \beta \in \mathbb{R}$$
Prove that affine functions are not linear. Prove that any affine function can be represented in the form $\mathbf{w}^T\mathbf{x} + \beta$.

12. Consider a function $f : \mathbb{R}^3 \to \mathbb{R}$, such that,
$$f\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = 2; \; f\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = -3; \; f\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = 1;$$
Can you determine the following values of $f(\mathbf{x})$, if you are told that f is linear?
$$f\left(\begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix}\right) =?; \; f\left(\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}\right) =?; \; f\left(\begin{bmatrix} 0.5 \\ 0.6 \\ -0.1 \end{bmatrix}\right) =?;$$
Can you find out these values if you are told that f is affine?

13. For the previous question, (a) assume that f is linear and find out $w \in \mathbb{R}^3$, such that $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x}$; and (b) assume f is affine and find out $\mathbf{w}, \beta$ such that $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + \beta$.

14. Consider the weighted norm of vector $\mathbf{v}$, defined as,
$$\|\mathbf{v}\|_{\mathbf{w}}^2 = \sum_{i=1}^{n} w_i v_i^2; \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$
Is this a valid norm?

15. Prove that the following modified version of the Cauchy-Bunyakovski-Schwartx Inequality is true.

$$\left|\sum_{i=1}^{n} u_i v_i w_i\right| \leq \|\mathbf{u}\|_{\mathbf{w}} \|\mathbf{v}\|_{\mathbf{w}}$$

16. Consider a basis $B = \{\mathbf{b}_i\}_{i=1}^{n}$ of R^n. Let the vector $\mathbf{x}$ with the following representations in the standard and B basis.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i \mathbf{e}_i \quad \text{and} \quad \mathbf{x}_b = \begin{bmatrix} x_{b1} \\ x_{b2} \\ \vdots \\ x_{bn} \end{bmatrix} = \sum_{i=1}^{n} x_{bi} \mathbf{b}_i$$

Evaluate the $\|\mathbf{x}\|_2^2$ and $\|\mathbf{x}_b\|_2^2$. Determined what happens to $\|\mathbf{x}_b\|_2^2$ under the following conditions on the basis vectors:

(a) $\|\mathbf{b}_i\| = 1, \forall i$

(b) $\left\|\mathbf{b}_i^T \mathbf{b}_j\right\| = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

17. Consider a set of measurements made from adult male subjects, where their height, weight and BMI (body moass index) were recorded as stored as vectors of length three; the first element is the height in cm, second is the weight in Kg, and the alst the the BMI. Consider the following four subjects,

$$\mathbf{s}_1 = \begin{bmatrix} 167 \\ 102 \\ 36.6 \end{bmatrix}; \ \mathbf{s}_2 = \begin{bmatrix} 180 \\ 87 \\ 26.9 \end{bmatrix}$$

$$\mathbf{s}_3 = \begin{bmatrix} 177 \\ 78 \\ 24.9 \end{bmatrix}; \ \mathbf{s}_4 = \begin{bmatrix} 152 \\ 76 \\ 32.9 \end{bmatrix}$$

You can use the distance between these vectors $\|\mathbf{s}_i - \mathbf{s}_j\|_2$ as a measure of the similiarity between the the four subjects. Generate a 4×4 table comparing the distance of each subject with respect to another subject; the diagonal elements of this table will be zero, and it will be symmetric about the main diagonal.

(a) Based on this table, how do the different subjects compare to each other?

(b) How do the similarities change if the height had been measured in m instead of cm? Can you explain this difference?

(c) Is there a way to fix this problem? Consider the weighted norm presented in one of the earlier problems.

$$\|x\|_{\mathbf{w}} = \left(w_1 x_1^2 + w_2 x_2^2 + \ldots + w_n x_n^2\right)^{\frac{1}{2}}$$

(d) What would be a good choice for $\mathbf{w}$ to address the problems with comparing distance between vectors due to change in units?

(e) Can the angle between two vectors be used as a measure of similarity between vectors? Does this suffer from the problem of $\|\mathbf{x}\|_2$?

Matrices

18. Elements of the matrix $\mathbf{C} \in \mathbb{R}^{m\times n}$ obtained as the product of two matrices $\mathbf{A} \in \mathbb{R}^{m\times p}$ and $\mathbf{B} \in \mathbb{R}^{p\times n}$ is given by,

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

We had discussed four different ways to think of matrix multiplication. By algebraically manipulating the previous equation arrive at these four views (inner product view, column view, row view and outer product view)?

19. Given the matrices $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & 3 & -1 & -1 \\ 2 & -2 & 0 & 2 \\ 1 & 0 & -3 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ -1 & 1 \\ -0 & 1 \end{bmatrix}$ and $\mathbf{C} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Evaluate the following products.

(a) $\mathbf{AB}$ (b) $\mathbf{A}^2\mathbf{B}$ (c) $\mathbf{CB}^T\mathbf{A}$ (d) $\mathbf{C}^3$ (e) $\mathbf{ABC}$

20. **Computational cost of different operations.** What is computational cost of the following matrix operations? Computational cost refers to the number of arithmetic operations required to carry out a particular matrix operation. Computational cost is a measure of the efficiency of an algorithm. For example, the consider the operation of vector addition, $\mathbf{a} + \mathbf{b}$, where $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. This requires n addition/subtraction operations and zero multiplication/division operations.

   (a) Matrix multiplication: $\mathbf{AB}$, where $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$
   (b) Inner product: $\mathbf{u}^T\mathbf{v}$
   (c) Gaussian Elimination for $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$.
   (d) Back substitution
   (e) Gauss-Jordan method to obtain the row echelon form: $\mathbf{A} \longrightarrow \mathbf{E}$.
   (f) Matrix inversion using the Gauss-Jordan method: $\left[\mathbf{A}|\mathbf{I}\right] \longrightarrow \left[\mathbf{I}|\mathbf{A}^{-1}\right]$

Report the counts for the addition/subtraction and multiplication/division operations separately.

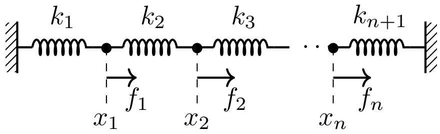
21. Prove $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

22. Consider the following matrix,

$$\mathbf{A} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 0.1 & 0 \\ 0 & 0.9 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

Find out the expression for $\mathbf{A}_n = \mathbf{A}^n$. What is $\mathbf{A}_\infty = \lim_{n\to\infty} \mathbf{A}^n$?

23. Derive force and displacement relationship for a series of $n+1$ springs (with spring constants k_i) connected in a line. There are n nodes, with f_i and x_i representing the force applied and resulting displacement at the i^{th} node.



   (a) Represent the relationship in the following form,

$$\mathbf{f} = \mathbf{Kx};\ \ \mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix};\ \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

   (b) What kind of a pattern does $\mathbf{K}$ have?
   (c) Consider a specific case where $n = 4$ and $k = 1.5N.m^{-1}$. What should be forces applied at the four nodes in order to displace the spring $\mathbf{x} = \begin{bmatrix} 0.5 \\ -0.5 \\ 0 \\ 0 \end{bmatrix} m$.

24. Prove that a matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$ can always be written as a sum a symmetric matrix $\mathbf{S}$ and a skew-symmetric matrix $\mathbf{A}$.

$$\mathbf{M} = \mathbf{S} + \mathbf{A},\ \ \mathbf{S}^T = \mathbf{S} \text{ and } \mathbf{A}^T = -\mathbf{A}$$

Does this property also hold for a complex matrix $\mathbf{M} \in \mathbb{C}^{n\times n}$?

25. The trace of a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ is defined as, $trace\,(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$. Prove the following,

   (a) $trace\,(\mathbf{A})$ is a linear function of $\mathbf{A}$.
   (b) $trace\,(\mathbf{AB}) = trace\,(\mathbf{BA})$
   (c) $trace\,(\mathbf{A}^T\mathbf{A}) = 0 \implies \mathbf{A} = 0$

26. Prove that the rank of an outer product $\mathbf{xy}^T$ is 1, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$.

27. Is there a relationship between the space of solutions to the following two equations?

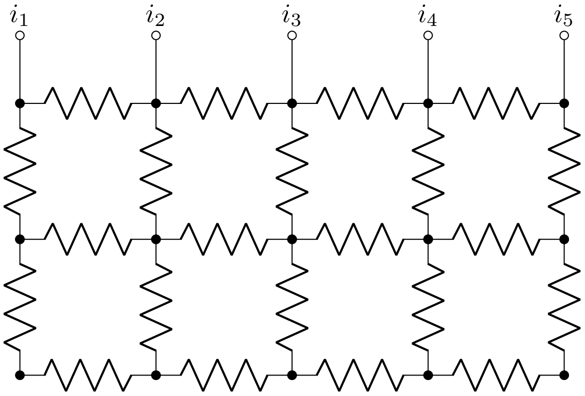
$$\mathbf{y}^T\mathbf{A} = \mathbf{c}^T \quad \text{and} \quad \mathbf{Ax} = \mathbf{b}$$

If so, how are they related?

28. Consider an upper triangular and lower triangular matrices $\mathbf{U}$ and $\mathbf{L}$, respectively.

   (a) Is the product of two upper triangular matrices $\mathbf{U}_1\mathbf{U}_2$ upper triangular?
   (b) Is the product of two lower triangular matrices $\mathbf{L}_1\mathbf{L}_2$ upper triangular?
   (c) What is the $trace\,(\mathbf{LU})$?

29. Consider the following electrical circuit with rectangular grid of resistors R. The input to this grid is a set of current injected at the top node as shown in the figure, such that $\sum_{k=1}^{5} i_k = 0$.



Express the relationship between the voltages at the different nodes (represented by • in the figure) and the net current flowing in/out of the node in the following form, $\mathbf{Gv} = \mathbf{i}$. Where, $\mathbf{G}$ is the conductance matrix, $\mathbf{v}$ is the vector of node voltages, and $\mathbf{i}$ is the vector representing the net current flow in/out of the different node.

30. Consider the following system.

$$\begin{bmatrix} 1 & 5 & 1 & 2 \\ 2 & -1 & 0 & 2 \\ 4 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \mathbf{x}_i = b_i$$

Solve the above equation using LU factorization for the following $\mathbf{b}_i$s.

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{b}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Construct a matrix $\mathbf{X}$ using the four solutions $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and $\mathbf{x}_4$ as its columns.

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \end{bmatrix}$$

Find out $\mathbf{XA}$ and $\mathbf{AX}$,. Based on this what can you say about $\mathbf{X}$?

31. How many different reduced row echelon forms can a matrix $\mathbf{A} \in \mathbb{R}^{4 \times 5}$ have? Hint: *Think in terms of basic and non-basic columns.*

32. Consider the system of equation, $\mathbf{Ax} = \mathbf{b}$, such that a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$. Are the following statements true? Explain your answer.

    (a) $rank\, \mathbf{A} \leq \min(m, n)$

    (b) The system is consistent if $rank\, \mathbf{A} = m$.

    (c) The system has a unique solution if $rank\, \mathbf{A} = n$.

33. Consider a linear function $f : V \to W$, where $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^m$. If V is a subspace of $\mathbb{R}^n$ then prove that W is a subspace of $\mathbb{R}^m$.

34. For a $n \times n$ square matrix $\mathbf{A}$, prove that if $\mathbf{AX} = \mathbf{I}$, then $\mathbf{XA} = \mathbf{I}$ and $\mathbf{X} = \mathbf{A}^{-1}$.

35. If two systems of linear equations are consistent, with augmented matices $[\mathbf{A}|\mathbf{b}]$ and $[\mathbf{A}|\mathbf{c}]$. Is $[\mathbf{A}|\mathbf{b} + \mathbf{c}]$ consistent?

36. Prove the following for the non-singular square matrices $\mathbf{A}$ and $\mathbf{B}$:

    (a) $\mathbf{AB}$ is non-singular.

    (b) $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$.

    (c) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

    (d) $\left(\mathbf{A}^T\right)^{-1} = \left(\mathbf{A}^{-1}\right)^T$

37. If a matrix $\mathbf{A}$ has LU decomposition, such that $\mathbf{A} = \mathbf{LU}$. Demonstrate that it also has a LDU decomposition $\mathbf{A} = \mathbf{LD\hat{U}}$, where $\mathbf{D}$ is a diagonal matrix, and $\hat{\mathbf{U}}$ is upper triangular. What happens to the LU and LDU decompositions when a matrix $\mathbf{A} = \mathbf{A}^T$?

38. Write down a basis for the four fundamental subspaces of the following matrix,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 & 4 & -1 & 0 \\ 4 & 8 & 12 & -8 & 2 & 1 \\ 2 & 3 & 2 & 1 & -2 & 0 \\ -3 & -1 & 1 & -4 & 0 & -1 \\ 1 & -2 & -1 & 0 & 0 & 0 \end{bmatrix}$$

39. Consider a matrix $\mathbf{A} = \begin{bmatrix} 1 & 4 & 5 \\ 4 & 118 & 26 \\ 3 & 16 & 30 \end{bmatrix}$.

    (a) Apply Gaussian elimination to simply this matrix into an upper-triangular matrix $\mathbf{U}$.

    (b) What is the corresponding upper-triangular matrix $\tilde{\mathbf{U}}$ obtained by applying Gaussian elimination to $\mathbf{A}^T$?

    (c) Could you have arrived at $\tilde{\mathbf{U}}$ without having to repeat the Gaussian elimination process on $\mathbf{A}^T$?

    (d) Write down the LDU decompositions of $\mathbf{A}$ and $\mathbf{A}^T$.

40. Derive the inverse of the matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

41. Consider the following upper-triangular matrix,

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

where, $u_{ii} \neq 0, \ 1 \leq i \leq n$. Do the columns of this matrix form a linearly independent set? Explain your answer.

42. Verify that $\mathbf{A}$ and $\mathbf{B}$ are inverses of each other,

    (a) $\mathbf{A} = \mathbf{I} - \mathbf{uv}^T$ and $\mathbf{B} = \mathbf{I} + \mathbf{uv}^T / \left(1 - \mathbf{v}^T\mathbf{u}\right)$

    (b) $\mathbf{A} = \mathbf{C} - \mathbf{uv}^T$ and $\mathbf{B} = \mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{uv}^T\mathbf{C}^{-1} / \left(1 - \mathbf{v}^T\mathbf{C}^{-1}\mathbf{u}\right)$

    (c) $\mathbf{A} = \mathbf{I} - \mathbf{UV}$ and $\mathbf{B} = \mathbf{I}_n + \mathbf{U}\left(\mathbf{I}_m - \mathbf{VU}\right)^{-1}\mathbf{V}$

(d) $\mathbf{A} = \mathbf{C} - \mathbf{U}\mathbf{D}^{-1}\mathbf{V}$ and $\mathbf{B} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1}$

where, $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{n}$, $\mathbf{U} \in \mathbb{R}^{n\times m}$, $\mathbf{V} \in \mathbb{R}^{m\times n}$ and $\mathbf{D} \in \mathbb{R}^{m\times m}$.

43. Consider the matrices $\mathbf{A} \in \mathbb{R}^{m\times m}$, $\mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{C} \in \mathbb{R}^{m\times n}$. Verify the following,

(a) $\begin{bmatrix}\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{B}\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{A}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{B}^{-1}\end{bmatrix}$

44. Gaussian elimination does not change the solution of a system $\mathbf{Ax} = \mathbf{b}$. Explain why the three row operations do not affect the solution of the system. Instead of row operations, what if we performed column operations. Will the solution of the system $\mathbf{Ax} = \mathbf{b}$ still remain unchanged? If the solution is affected, how is it affected by the following operations?

(a) Columns $\mathbf{a}_i$ and $\mathbf{a}_j$ of $\mathbf{A}$ are interchanged.

(b) Column $\mathbf{a}_i$ is replaced by $\alpha\mathbf{a}_i$.

(c) Columns $\mathbf{a}_i$ is replace by $\mathbf{a}_i + \beta\mathbf{a}_j$.

45. **Two point boundary problem.** $\mathbf{Ax} = \mathbf{b}$ is often encountered in many practical applications. One such application is the numerical solution of differential equations of the following form,

$$\sum_{i=0}^{M} a_i\left(x\right) y^{(i)}\left(x\right) = f\left(x\right)$$

where, $x \in [a, b]$ and $y\left(a\right) = \alpha, y\left(b\right) = \beta$.

Numerical methods are often employed for obtaining an approximate estimate of $y\left(x\right)$ at discrete points in the interval $[a, b]$. The interval is divided into subintervals of width Δx. The derivate of $y\left(x\right)$ at the different nodes (points between two subintervals) can be approximated as the following,

$$y'\left(x_i\right) = \frac{y\left(x_i + \Delta x\right) - y\left(x_i - \Delta x\right)}{\Delta x}$$
$$y''\left(x_i\right) = \frac{y\left(x_i + \Delta x\right) + 2y\left(x_i\right) - y\left(x_i - \Delta x\right)}{\Delta x^2}$$

where, $x_i = a + i\Delta x,\ 0 \leq i \leq N+1$, and $b - a = \left(N+1\right)\Delta x$. Addition and subtracting the above two equations and neglecting terms involving higher orders of Δx, we get the following approximations for the derivatives of $y\left(x\right)$ at x_i.

Replacing the derivatives of $y\left(x\right)$ by the above approximations and evaluating the equation at the different nodes x_is, we arrive a set of N linear equations with N unknowns $y\left(x_1\right), y\left(x_2\right), \ldots y\left(x_N\right)$.

Using this approach, compute an approximate solution for $y\left(x\right)$ for the following differential equations over the interval $x \in [0, 1]$.

(a) $y''\left(x\right) = -x$

(b) $y''\left(x\right) + y'\left(x\right) = x$

Solve these equations for different values of Δx, and compare the resulting approximate solution for $y\left(x\right)$ with the exact solution. Present your results as a plot the solution $y\left(x_i\right)$ versus x_i.

Comment on the dependence of the solution $\left(x\right)$ on Δx. What is the best value for Δx to use in solving these equations?

46. **Ill-conditioned systems.** A system $\mathbf{Ax} = \mathbf{b}$ is said to be ill-conditioned when small changes in the components of $\mathbf{A}$ or $\mathbf{b}$ can produce large changes in the solution $\mathbf{x}$. Consider the following system,

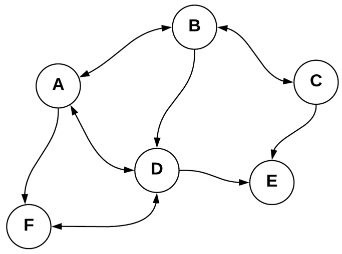
$$x - y = 100$$
$$10 + \left(9 + \Delta\right) y = 0$$

Find the solutions of the system for different values of $\Delta = -2, -1, 0, 1, 2$. How do the solutions change with Δ. Now consider the following system,

$$x - y = 100$$
$$10 - \left(9 + \Delta\right) y = 0$$

The second system is an example of an ill-conditioned system. What can you say about the geometries of these two systems?

47. **Connectivity matrices.** Another common application of matrices is in graph theory. A graph is a set of vertices or nodes connected by edges, as show in the following figure. A-F are the nodes of the graph, and the lines with the arrows are the edges that convey information about the connections or relationships between the nodes.



The above graph can be thought as a representation of different places in a city (represented by the nodes), and the lines with the arrows represent the roads connecting these different places. A line with two arrows allow two-way traffic, while line with single arrow only allow one way traffic. The connectivity between the different places can be summarized though the connectivity matrix $\mathbf{C} \in \mathbb{R}^{n\times n}$, where n is the number of nodes in the graph. The elements of this connectivity matrix represents whether or not there is a direct path between two places.

$$c_{ij} = \begin{cases} 1 & \text{there is a direct road between places } i\,\&\,j. \\ 0 & \text{otherwise.}\end{cases}$$

The diagonal element of $\mathbf{C}$ are zero, $c_{ii} = 0$.

Write down the connectivity matrix $\mathbf{C}$ for the graph shown above. How can we use the matrix $\mathbf{C}$ to answer the following questions? Explain exact matrix operation you would perform to answer these questions (Hint: Consider higher power of $\mathbf{C}$).

(a) Is there a path between two places i and j that goes via one other place? For example, we can go from A to D via B.

(b) How many paths are there between places i and j that goes via three other places?

Orthogonality

48. Consider an orthonormal set of vectors $V = \{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r\}, \ \mathbf{v}_i \in \mathbb{R}^n \ \forall i \in \{1, 2, \dots r\}$. If there is a vector $\mathbf{w} \in \mathbb{R}^n$ such that $\mathbf{v}_i^T \mathbf{w} = 0 \ \forall i \in \{1, 2, \dots r\}$. Prove that $\mathbf{w} \notin span\,(V)$.

49. Consider the following set of vectors in $\mathbb{R}^4$.
$$V = \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \\ 4 \end{bmatrix} \right\}$$
Find the set of all vectors that are orthogonal to V?

50. For a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, prove that $C(\mathbf{A}) \perp N\left(\mathbf{A}^T\right)$ and $C\left(\mathbf{A}^T\right) \perp N(\mathbf{A})$.

51. If the columns of a matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ are orthonormal, prove that $\mathbf{A}^{-1} = \mathbf{A}^T$. What is $\mathbf{A}^T\mathbf{A}$ when $\mathbf{A}$ is rectangular $(\mathbf{A} \in \mathbb{R}^{m \times n})$ with orthonormal columns?

52. What will happen when the Gram-Schmidt procedure is applied to: (a) orthonormal set of vectors; and (b) orthogonal set of vectors? If the set of vectors are columns of a matrix $\mathbf{A}$, then what are the corresponding $\mathbf{Q}$ and $\mathbf{R}$ matrices for the orthonormal and orthogonal cases?

53. Consider the linear map, $\mathbf{y} = \mathbf{A}\mathbf{x}$, such that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n \times n}$. Let us assume that $\mathbf{A}$ is full rank. What conditions must $\mathbf{A}$ satisfy for the following statements to be true,
    (a) $\|\mathbf{y}\|_2 = \|\mathbf{x}\|_2$, for all $\mathbf{x}, \mathbf{y}$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$.
    (b) $\mathbf{y}_1^T \mathbf{y}_2 = \mathbf{x}_1^T \mathbf{x}_2$, for all $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2$ such that $\mathbf{y}_1 = \mathbf{A}\mathbf{x}_1$ and $\mathbf{y}_2 = \mathbf{A}\mathbf{x}_2$.

    Note: A linear map $\mathbf{A}$ with the aforementioned properties preserves lengths and angle between vectors. Such maps are encountered in rigid body mechanics.

54. Prove that the rank of an orthogonal projection matrix $\mathbf{P}_S = \mathbf{U}\mathbf{U}^T$ onto a subspace $\mathcal{S}$ is equal to the dim $\mathcal{S}$, where the columns of $\mathbf{U}$ form an orthonormal basis of $\mathcal{S}$.

55. If the columns of $\mathbf{A} \in \mathbb{R}^{m \times n}$ represent a basis for the subspace $\mathcal{S} \subset \mathbb{R}^m$. Find the orthogonal projection matrix $\mathbf{P}_{\mathcal{S}}$ onto the subspace $\mathcal{S}$. Hint: Gram-Schmidt orthogonalization.

56. Consider two orthogornal matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$. Is the $\mathbf{Q}_2^T\mathbf{Q}_1$ an orthogonal matrix? If yes, prove that it is so, else provide a counter-example showing $\mathbf{Q}_2^T\mathbf{Q}_1$ is not orthogonal.

57. Let $\mathbf{P}_{\mathcal{S}}$ represent an orthogonal projection matrix onto to the subspace $\mathcal{S} \subset \mathbb{R}^n$. What can you say about the rank of the matrix $\mathbf{P}_{\mathcal{S}}$? Explain how you can obtain an orthonormal basis for $\mathcal{S}$ from $\mathbf{P}_{\mathcal{S}}$.

58. Consider a 1 dimensional subspace spanned by the vector $\mathbf{u} \in \mathbb{R}^n$. What kind of a geometric operation does the matrix $\mathbf{I} - 2\frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T\mathbf{u}}$ represent?

59. Prove that when a triangular matrix is orthogonal, it is diagonal.

60. If an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is to be partitioned such that, $\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix}$, then prove that $C(\mathbf{Q}_1) \perp C(\mathbf{Q}_2)$.

61. Find an orthonormal basis for the subspace spanned by $\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 3 \end{bmatrix} \right\}$.

Matrix Inverses

62. Consider the following bases for $\mathbb{R}^3$.

$$A^S = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$$

$$B^A = \left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Where, X^Y is the basis X represented in another basis Y; S stands for the standard basis. Let $\mathbf{b}_X$ stand for the representation of vector in $\mathbb{R}^3$ in the basis X.

(a) Consider a vector $\mathbf{b}_S = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$ represented in the standard basis. What is the representation of $\mathbf{b}_S$ in the other four basis A, and B?

(b) Consider a vector $\mathbf{d}_B = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$ represented in the basis B. What is the representation of this vector in the standard basis?

63. When does the following diagnoal matrix have an inverse?

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}$$

Write down an expression for D^{-1}.

64. Prove that the inverse of a non-singular upper-triangular matrix is upper-triangular. Using this show that for a lower triangular matrix it is lower-triangular.

65. Consider a 2×2 block matrix, $\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix}$, where $\mathbf{A} \in \mathbb{R}^{m \times m}$. Find an expression for the inverse $\mathbf{A}^{-1}$ interms of the block components and their inverses (if they exist) of $\mathbf{A}$. Hint: Consider $\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix}$, and solve $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

66. Express the inverse of the following matrix in terms of $\mathbf{A}$ and $\mathbf{b}$.

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0} & 1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

where, $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{b} \in \mathbb{R}^n$.

67. Consider a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with linearly independent columns. Prove that the Gram matrix $\mathbf{A}^T\mathbf{A}$ is invertible.

68. Find all possible left/right inverses for the following matrices, if they exist.

(a) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -2 & -3 & -4 \end{bmatrix}$

(b) $\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$

(c) $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ -3 & 4 \end{bmatrix}$

(d) $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$

For each of these matrices find the corresponding pseudo-inverse $\mathbf{A}^{\dagger}$, and verify that the pseudo-inverse has the minimum squared sum of its components.

69. Prove that the inverse of a non-singular symmetric matrix is symmetric.

70. Consider the scalar equation, $ax = ay$. Here we can cancel a from the equation when $a \neq 0$. When can we carry out similar cancellations for matrcies?

(a) $\mathbf{AX} = \mathbf{AY}$. Prove that here $\mathbf{X} = \mathbf{Y}$ only when $\mathbf{A}$ is left invertible.

(b) $\mathbf{XA} = \mathbf{YA}$. Prove that here $\mathbf{X} = \mathbf{Y}$ only when $\mathbf{A}$ is right invertible.

71. Consider two non-singular matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$. Explain whether or not the following matrices are invertible. If they are, then provide an expression for it inverse.

(a) $\mathbf{C} = \mathbf{A} + \mathbf{B}$

(b) $\mathbf{C} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$

(c) $\mathbf{C} = \begin{bmatrix} \mathbf{A} & \mathbf{A} + \mathbf{B} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$

(d) $\mathbf{C} = \mathbf{ABA}$

72. Consider the matrices $\mathbf{A} \in \mathbb{R}^{m \times l_1}$ and $\mathbf{B} \in \mathbb{R}^{l_2 \times m}$. Can you find the requirements for matrices $\mathbf{A}$ and $\mathbf{B}$, such that $\mathbf{AXB} = \mathbf{I}$, where $\mathbf{X} \in \mathbb{R}^{l_1 \times l_2}$? Assuming those conditions are satisfied, find an expression for $\mathbf{X}$?

73. Consider a matrix $\mathbf{C} = \mathbf{AB}$, where $\mathbf{A} \in \mathbf{R}^{m \times n}$ and $\mathbf{B} \in \mathbf{R}^{n \times m}$. Explain why $\mathbf{C}$ is not invertible when $m > n$. Suppose $m < n$, under what conditions is $\mathbf{C}$ invertible?

74. For a square matrix $\mathbf{A}$ with non-signular $\mathbf{I} - \mathbf{A}$, prove that $\mathbf{A}\left(\mathbf{I} - \mathbf{A}\right)^{-1} = \left(\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{A}$.

75. Consider the non-singular matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{A} + \mathbf{B}$. Prove that,

$$\mathbf{A}\left(\mathbf{A} + \mathbf{B}\right)^{-1}\mathbf{B} = \mathbf{B}\left(\mathbf{A} + \mathbf{B}\right)^{-1}\mathbf{A} = \left(\mathbf{A}^{-1} + \mathbf{B}^{-1}\right)^{-1}$$

Eigenvalues and Eigenvectors

76. Explain why an eigenvector cannot be associated with two eigenvalues.

77. What are the eigenspaces associated with the diagonal matrix $\mathbf{D} = \text{diag}\,(d_1, d_2, \ldots d_n)$?

78. If a matrix $\mathbf{A}$ has zero as one of its eigenvalues, explain why $\mathbf{A}$ must be singular.

79. For a matrix $\mathbf{A}$ with eigenvalues $\{\lambda_i\}_{i=1}^n$, verify for the following matrices that $\Pi_{i=1}^n \lambda_i = \det(\mathbf{A})$ and $\sum_{i=1}^n \lambda_i = trace\,(\mathbf{A})$.

    (a) $\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$

    (b) $\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}$

    (c) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

    (d) $\frac{1}{5}\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \end{bmatrix}$

80. Let $\{\lambda_i, \mathbf{v}_i\}_{i=1}^n$ be the eigenpairs of a matrix $\mathbf{A}$. Then prove that,

    (a) $\left\{\lambda_i^k, \mathbf{v}_i\right\}_{i=1}^n$ are the eigenpairs of $\mathbf{A}^k$.

    (b) $\{p(\lambda_i), \mathbf{v}_i\}_{i=1}^n$ are the eigenpairs of $p(\mathbf{A})$, where $p(\mathbf{A}) = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} + \ldots + \alpha_k\mathbf{A}^k$.

81. Prove that if $\{\lambda_i, \mathbf{v}_i\}_{i=1}^n$ are the eigenpairs of a matrix $\mathbf{A}$, then the eigenpairs of $\mathbf{A}^k$ are $\left\{\lambda_i^k, \mathbf{v}_i\right\}_{i=1}^n$.

82. Consider the matrices $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$. Are the eigenvalues of $\mathbf{AB}$ equal the eigenvalues of $\mathbf{BA}$?

83. Consider the matrices $\mathbf{A}$ and $\mathbf{B}$. If $\mathbf{v}$ is an eigenvector $\mathbf{B}$, underwhat condition will $\mathbf{v}$ also be the eignevector of $\mathbf{AB}$. Under these conditions, what will be corresponding eigenvalue of $\mathbf{v}$? How do your answers change in the case of $\mathbf{BA}$?

84. Let $\{\lambda_i, \mathbf{v}_i\}_{i=1}^n$ are the eignepairs of a matrix $\mathbf{A}$. What are the eigenpairs of the folliowing?

    (a) $2\mathbf{A}$

    (b) $\mathbf{A} - 2\mathbf{I}$

    (c) $\mathbf{I} - \mathbf{A}$

85. Let $\mathbf{A} = \begin{bmatrix} 0.6 & 0.2 \\ 0.4 & 0.8 \end{bmatrix}$. What is the value of: (a) A^2 (b) A^{100} (c) A^{∞}?

86. Show that $\mathbf{u} \in \mathbb{R}^2$ is an eigenvector of $\mathbf{A} = \mathbf{u}\mathbf{v}^T$. What are the two eigenvalues of $\mathbf{A}$?

87. Consider two similar matrices $\mathbf{A}$ and $\mathbf{B}$. Prove that the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$ are the same. How are the eigenvectors of $\mathbf{A}$ and $\mathbf{B}$ related to each of other for a given eigenvalue?

88. Find the eigenvectors of the following permutation matrix $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

89. **Left eigenvectors**: Consider a matrix $\mathbf{A}$ with eigenpairs $\{\lambda_1, \mathbf{v}_i\}_{i=1}^n$. The left eigenvectors of the matrix $\mathbf{A}$ are the vectors that satisfy the equation, $\mathbf{A}^T\mathbf{w} = \mu\mathbf{w}$ (or $\mathbf{w}^T\mathbf{A} = \mu\mathbf{w}^T$), and let $\{\mu_i, \mathbf{w}_i\}_{i=1}^n$ be the left eigenpairs of $\mathbf{A}$. Show the following,

    (a) The eigenvalues of both $\mathbf{A}$ and $\mathbf{A}^T$ are the same.

    (b) $\mathbf{v}_i^T\mathbf{w}_j = 0$. The eigenvector $\mathbf{v}_i$ corresponding to the eigenvalue λ_i and the left eigvenvector $\mathbf{w}_j$ corresponding to the eigenvalue λ_j are orthogonal, when $\lambda_i \neq \lambda_j$.

    (c) The matrix A can be expressed as a sum of rank-one matrices,

    $$\mathbf{A} = \lambda_1\mathbf{v}_1\mathbf{w}_1^T + \lambda_2\mathbf{v}_2\mathbf{w}_2^T + \ldots + \lambda_n\mathbf{v}_n\mathbf{w}_n^T$$

90. Prove that $\mathbf{AA}^T$ has real and positive eigenvalues, and that the eigenvectors corresponding to distinct eigenvalues of $\mathbf{AA}^T$ are orthogonal.

91. If $\{\lambda_i, \mathbf{v}_i\}_{i=1}^n$ are the eigenpairs of a non-singular matrix $\mathbf{A}$, the prove that $\left\{\lambda_i^{-1}, \mathbf{v}_i\right\}_{i=1}^n$ are the eigenpairs of $\mathbf{A}^{-1}$.

92. A matrix $\mathbf{A}$ is called *nilpotent* if $\mathbf{A}^k = \mathbf{0}$ for some finite positive integer k. Prove that the $trace\,(\mathbf{A}) = 0$ for a nilpotent matrix $\mathbf{A}$. What are all the eigenvalues of such a matrix?

Positive Definite Matrices and Matrix Norm

93. Prove that $\mathbf{A}^T\mathbf{A}$ is positive semi-definite for any matrix $\mathbf{A}$. When is $\mathbf{A}^T\mathbf{A}$ guaranteed to be positive definite?

94. If $\mathbf{A}$ is positive definite, then prove that $\mathbf{A}^{-1}$ is also positive definite.

95. Show that a positive definite matrix cannot have a zero or a negative element along its diagonal.

96. Show that the following statements are true.

    (a) All positive definite matrices are inverstible.

    (b) The only positive definite projection matrix is $\mathbf{I}$.

97. Is the function $f(x_1, x_2, x_3) = 12x_1^2 + x_2^2 + 6x_3^2 + x_1x_2 - 2x_2x_3 + 4x_3x_1$ positive definite?

98. The LU decomposition for symmetric matrices can be written as $\mathbf{A} = \mathbf{L}^T\mathbf{D}\mathbf{L}$, where $\mathbf{D}$ is a diagonal matrix, and $\mathbf{L}$ is lower triangular with 1 along its main diagonal. When $\mathbf{A}$ is postive definite, we can write, $\mathbf{A} = \mathbf{C}^T\mathbf{C} = \mathbf{L}^T\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}\mathbf{L}$. This is the *Cholesky decomposition*. Find $\mathbf{C}$ for the following,

    (a) $\begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$

    (b) $\begin{bmatrix} 4 & 1 & 0 \\ 1 & 8 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

99. Prove the following for $\mathbf{A} \in \mathbb{R}^{m \times n}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{a}_1^T} \\ \tilde{\mathbf{a}_1^T} \\ \vdots \\ \tilde{\mathbf{a}_m^T} \end{bmatrix}$$

    (a) $\|\mathbf{A}\|_1 = \max_{1 \leq i \leq n} \|\mathbf{a}_i\|_1$

    (b) $\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq m} \|\tilde{\mathbf{a}}_i\|_1$

    (c) $\|\mathbf{A}\|_2 = \max_{1 \leq i \leq n} |\lambda_i|$, where λ_i are the eigenvalues of $\mathbf{A}^T\mathbf{A}$.

    (d) $\|\mathbf{A}\|_F = trace\left(\mathbf{A}^T\mathbf{A}\right)$

100. Prove that the induced norm of a matrix product is bounded: $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|$.

101. Verify the following inequalities on vector and matrix norms ($\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$):

    (a) $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2$

    (b) $\|\mathbf{x}\|_2 \leq \sqrt{m}\|\mathbf{x}\|_\infty$

    (c) $\|\mathbf{A}\|_\infty \leq \sqrt{n}\|\mathbf{A}\|_2$

    (d) $\|\mathbf{A}\|_2 \leq \sqrt{m}\|\mathbf{A}\|_\infty$

102. Find an expression for the induced 2-norm of an outer product, $\mathbf{A} = \mathbf{u}\mathbf{v}^T$, where $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$.

Singular Value Decomposition

103. For a square $\mathbf{A} \in \mathbb{R}^{n\times n}$, the SVD tells us how a unit sphere in $\mathbb{R}^n$ is distorted by the linear transformation performed by $\mathbf{A}$. This degree of distortion can be quantified using the singular values of $\mathbf{A}$, which is the 2-norm *condition number*,

$$\kappa = \frac{\sigma_1}{\sigma_n}$$

   (a) Explain why $\kappa \geq 1$?

   (b) What is condition number of a singular matrix?

   (c) If $\mathbf{A}$ is non-singular, show that $\kappa = \|\mathbf{A}\|_2 \left\|\mathbf{A}^{-1}\right\|_2$

   (d) Condition numbers can also be defined based on other p-norms. The general p-norm condition number is given by, $\kappa_p = \|\mathbf{A}\|_p \left\|\mathbf{A}^{-1}\right\|_p$. Evaluate the 1-norm, 2-norm and ∞-norm condition numbers for the following matrices. How do these number compare with each other? (i) $\mathbf{A} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$; (ii) $\mathbf{A} = \begin{bmatrix}1 & -1\\10 & -9\end{bmatrix}$; (iii) $\mathbf{A} = \begin{bmatrix}1 & 5\\-1 & 1\end{bmatrix}$.

   (e) Conditions numbers play an important role in practice. We had earlier an example of an ill-conditioned system $\mathbf{Ax} = \mathbf{b}$ (problem 46). Consider the following systems, where: (i) $\mathbf{A}_1 = \begin{bmatrix}1 & -1\\10 & -9\end{bmatrix}$; and (ii) $\mathbf{A}_2 = \begin{bmatrix}1 & -10\\1 & 10\end{bmatrix}$.

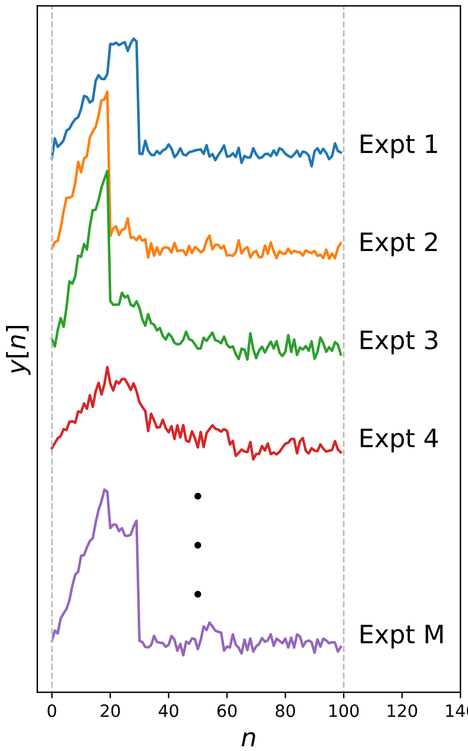
   For $\mathbf{b} = \begin{bmatrix}10\\0\end{bmatrix}$, what are the solutions $\mathbf{x}_1 \left(= \mathbf{A}_1^{-1}\mathbf{b}\right)$ and $\mathbf{x}_2 \left(= \mathbf{A}_2^{-1}\mathbf{b}\right)$?

   Suppose there is an error in the measurement of $\mathbf{b}$, and we have $\tilde{\mathbf{b}} = \begin{bmatrix}9\\1\end{bmatrix}$. The relative error in $\mathbf{b}$ is given by $\delta b = \frac{\|\mathbf{b}-\tilde{\mathbf{b}}\|_2}{\|\mathbf{b}\|_2}$. What are the new solutions $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{x}}_2$?

   Calculate δx_1 and δx_2, the relative errors in $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively? How do these compare to δb?

   Note: Through this problem, you should be able to see that an ill-conditioned system has a large condition number, which can amplify error and thus lead to large uncertainty in the solutions.

104. Consider a system S, which when probed with a test signal $x(t)$, and the system responds with an output signal $y(t)$. The experiment is repeated M times on the system by repeatedly applying the text signal to th system and the response of the system is sampled $y[n]$ and recorded for a for a fixed duration of time $0 \leq n \leq N$. An example of the response of the system for the different experiments are shown below $(N = 100)$,



We know from the physics of the system that the system response to the test signal for the j^{th} experiment given is given by ,

$$y_j[n] = \sum_{i=1}^{K} w_{ji}\phi_i[n] + \nu_j[n]$$

where, n is time index; N is the number of data points recorded on each experiment; $\phi_i[n], 1 \leq i \leq K$ are characteristic signals of the system; w_{ji} are the weights that determine the amount of the ϕ_i present in the output y_j for the j^{th} experiment; and $\nu_j[n]$ is measurement noise present in the j^{th} experiments. Additionally, we also know that

$$\phi_i^T\phi_j = \sum_{l=0}^{N-1}\phi_i[l]\,\phi_j[l] = \begin{cases}0 & i \neq j\\ 1 & i = j\end{cases}$$

Data from one such experiment is provided in the CSV file (trigexpt.csv) which contains an array of data where the rows correspond to system response for the different experiments, and the columns correspond to different time index. Using this data identify the different characteristic signals ϕ_i for the system, along with the weights w_{ji} for the different experiments. You must also ensure that the data is explained with least number characteristic signal, i.e. K should be as low as possible. Explain how you chose K.l

Least Squares

105. The least square approximate solution to problem $\mathbf{Ax} = b, \mathbf{A} \in \mathbb{R}^{m\times n}$ is obtained by minimizing the following objective fucntion,

$$O(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|^2 = \sum_{i=1}^{m} \left(\tilde{\mathbf{a}}_i^T \mathbf{x} - \mathbf{b}_i\right)^2$$

If the the objective function was defined differently, where the different components were given different weights for the different terms,

$$O_w(\mathbf{x}) = \sum_{i=1}^{m} w_i \left(\tilde{\mathbf{a}}_i^T \mathbf{x} - \mathbf{b}_i\right)^2$$

This is the *weighted least squares*. Find the expression for the approximate solution of the weighted least squares problem.

106. Consider a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ with independent columns. There is no matrix $\mathbf{X}$, such that $\mathbf{AX} = \mathbf{I}$. Find an expression for the matrix $\mathbf{X}$, such that $\|\mathbf{AX} - \mathbf{I}\|^2$ is minimum. Hint: Consider the individual columns of $\mathbf{X}$ and $\mathbf{I}$.

107. Consider the following polynomial equation,

$$y = \sum_{i=0}^{n} \beta_i x^i, \quad x, y, \beta_i \in \mathbb{R}$$

This expression is linear in the polynomial coefficients. Fitting a polynomial to data can be done through a linear least square procedure. Conisder a set of measurements $\{(x_l, y_l)\}_{l=1}^{m}$. We are interested in fitting a polynomial that fits this data, such that the difference between the polynomial and data is as low as possible.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\beta$$

$\mathbf{X}$ is called the *Vandermonde* matrix. To estimate β through a least squares procedure, $\mathbf{X}$ must have independent columns. Prove that for a set of different x_is the *Vandermonde* matrix has independent columns.

You are provided with CSV format data file (polyfit.csv) containing data $\{(x_l, y_l)\}_{l=1}^{m}$, to which you are required to fit a polynomial function. The choice of the order of the polynomial is up to you. This can be come from *a priori* knowledge of the system from which data is collected, or it can be decided based on eyeballing the scatter plot between x_i and y_i.

(a) **Data fitting error**: Fit polynomials of orders 0 to 10 to the given data, and for each order determine the data fitting error.

$$e_n = \left\|\mathbf{X}\hat{\beta} - \mathbf{y}\right\|$$

Plot e_n versus the polynomial model order n. You should observe the fitting error to monotonically decrease as a function fo n. Why is this? Can e_n ever be zero?

(b) **Model validation**: Even though increasing the polynomial order decreases the data fitting error, this is not always desirable. Given that noise in ubiquitous in all measurements, increasing model order will result in a polynomial that not only fits the general trend in the data, but also the observed measurement noise. Thus, the optimal choice for the model order is determined through a validation procedure, where the data is split into two sets - *training* set and a *testing* set. In order to understand this, you are required to:

   i. Split your data D into two sets of size 80% and 20%, corresponding to the *training* D_{train} and *testing* D_{test} sets; these percentages are arbitrary. Split you data randomly, such that each data entry is randomly assigned to D_{train} and D_{test}.

   ii. Fit the polynomial model of a particular order to D_{train}. Let the model parameters obtained be $\hat{\beta}$. The validation error for this model is defined as the following.

$$e_n^{val} = \left\|\mathbf{X}\hat{\beta} - \mathbf{y}\right\|$$

   where, $\mathbf{X}$ and $\mathbf{y}$ comes from the test data set D_{test}. Estimate the validation error for different models order 0 to 10, and plot e_n^{val} versus the polynomial model order n. How is this plot different from the plot e_n versus n? What is the optimal choice for the model order based on the validation procedure?

(c) **Regularized data fitting**: Instead of minimzing $\|\mathbf{X}\beta - \mathbf{y}\|^2$ of the data, now fit a model that minimizes, $\|\mathbf{X}\beta - \mathbf{y}\|^2 + \lambda\beta^T\beta$, where $\lambda \geq 0$. In this particular case fit the model order to a high value (e.g. 10) and the entire data set D. Perform the data ditting procedure for different values of λ. Plot $\|\mathbf{X}\beta - \mathbf{y}\|$ verus λ. Compare the values of $\hat{\beta}$ for the different values of λ and compare these to your optimal choice of model parameters from the previous question.

108. Consider a time series $\mathbf{x} = \{x_0, x_1, \dots x_{N-1}\}$ consisting of N data points, where n indicates time index. The time series is corrupted by noise, and we are interested in filtering the time series to obtain a smooth estimate of the the general trend in the time series. This can be posed a problem of estimating a new time series $\hat{\mathbf{x}}$, such that the difference $\mathbf{x} - \hat{\mathbf{x}}$ is minimized and $\hat{\mathbf{x}}$ is smooth, i.e. the adjacent values of the signal do not change abruptly.

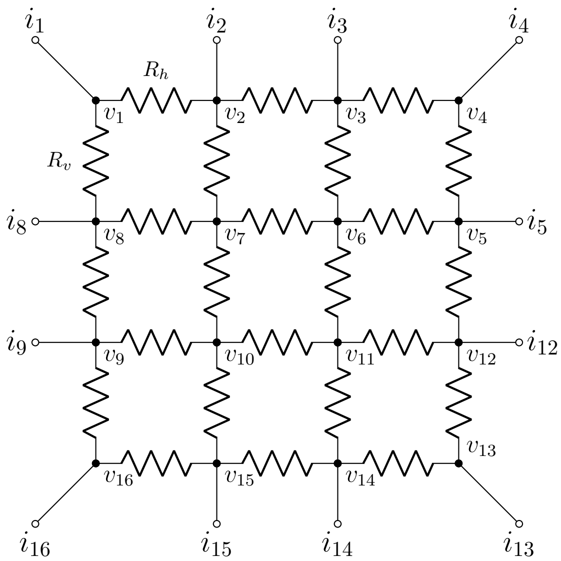
$$O(\hat{x}) = \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 + \lambda \sum_{i=2}^{N-1} (2\hat{x}_i - \hat{x}_{i-1} - \hat{x}_{i+1})^2$$

If $\mathbf{x}$ and $\hat{\mathbf{x}}$ are considered as N-vectors, then this can eb written as

$$O\left(\hat{\mathbf{x}}\right) = \left\|\mathbf{x} - \hat{\mathbf{x}}\right\|^2 + \lambda \left\|\mathbf{D}\hat{\mathbf{x}}\right\|^2$$

You are provided with a CSV data file (timeseries.csv) consisting of a time series. Filter this time series by minimizing $O\left(\hat{\mathbf{x}}\right)$ for different values of λ. Plot $\mathbf{x}$ and $\hat{\mathbf{x}}$ for different values of λ. What role does λ play in the minimzation problem?

109. Consider the following resistive network, where the horizontal resistors have resistnce of $R_h = 1\Omega$ and the vertical resistors have a resistance of $R_v = 2\Omega$. You goal is to determine a set of currents in the $\mathbf{i} = \begin{bmatrix} i_1 & i_2 & \dots & i_{16} \end{bmatrix}^T$, so as to acheive a particular distribution of potentials at the different nodes of the network. Node that we have control only over the currents at the edge nodes, and in the internal nodes the current is zero, i.e. $i_6 = i_7 = i_{10} = i_{11} = 0$.



Let $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_{16} \end{bmatrix}$ represent the vector of potential distributions in the network. Then determine $\mathbf{i}$ such that $\left\|\mathbf{v}_T - \mathbf{v}\right\|^2$ is minimized for the followingdesired potential distribution (Note that the potentias are arragned in a matrix $\mathbf{V}_{map} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \\ v_8 & v_7 & v_6 & v_5 \\ v_9 & v_{10} & v_{11} & v_{12} \\ v_{16} & v_{15} & v_{14} & v_{13} \end{bmatrix}$

(a) $\mathbf{V}_{map} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$

(b) $\mathbf{V}_{map} = \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & 1 & 1 & 2 \\ 2 & 1 & 1 & 2 \\ 2 & 2 & 2 & 2 \end{bmatrix}$

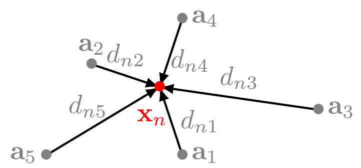
(c) $\mathbf{V}_{map} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

for some desired potential distribution $\mathbf{v}_T$, subject to the constraint $\sum_{k=1}^{16} i_k = 0$.

110. Trialteration is a process of determining the position $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ of a point P given the distance of the point from a N control points of known locations $\mathbf{a}_1, \mathbf{a}_2, \dots \mathbf{a}_N$. At any given point in time n, we have N different measurements corresponding to the distance of the point P from the N control points, i.e.

$$\mathbf{d}_n = \begin{bmatrix} d_{n1} & d_{n2} & \dots & d_{nN} \end{bmatrix}^T$$

where, $d_{ni}^2 = \left\|\mathbf{x}_n - \mathbf{a}_i\right\|_2^2$ for all $1 \leq i \leq N$.



$\mathbf{x}_n$ are nonlinear functions of the measurements and the control point locations. However, taking the difference between two squared distance measurements d_j^2 and d_i^2 results in linear equations in the unknown coordinates,

$$\left(\mathbf{a}_i - \mathbf{a}_j\right)^T \mathbf{x} = \frac{\left(d_j^2 - \left\|\mathbf{a}_j\right\|_2^2\right) - \left(d_i^2 - \left\|\mathbf{a}_i\right\|_2^2\right)}{2}$$

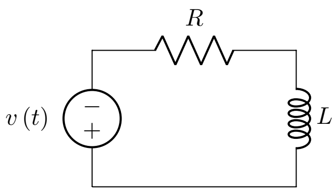
Consider the point P that moves with time; $\mathbf{x}_n \in \mathbb{R}^2$ represents the location of the point at time instant n. The distance of this point from a set of 50 different control points is given to you in a CSV data file (trilatdist.csv); the data rows correspond to distance measurements from the different control points at a given point in time, while the colums correpond to the distance of the point from a particular control point for all time. The locations of the 50 control points are available in a separate fileCSV file (trilatctrlpos.csv). Each of these distance measurements is affected by noise. You aim is to use these distance measurements to reconstruct the trajecotry of the point $\mathbf{x}_n$ as a function of time.

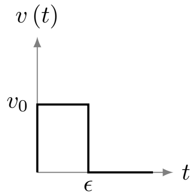
(a) What is the minimum number of distance measurements you need to estimate the unknown coordinates of the point $\mathbf{x}$?

(b) You are provided the actual trajectory of point P in the file trilatactpos.csv. How does you estimate compare with that of the actual trajectory? If we are informed that the point P does not undergo large changes in position between two consecutive time instants $n-1$ and n, how would you use this information to improve your estimate?

Linear Dynamics Systems – Transfer function view

111. Write down the differential equation representing relationship between the loop current $y(t)$ and the input voltage $v(t)$. Assume a initial loop current of $y(0^-) = y_0$.





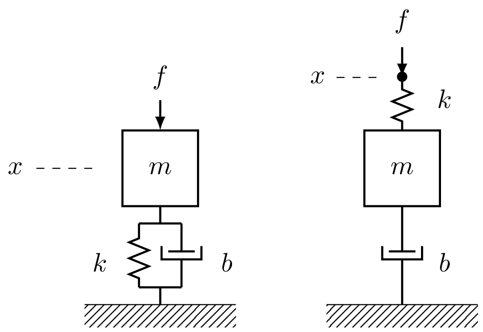
(a) Find the response of the system for the input $v(\bullet), \forall t \geq 0$ shown in the figure.

(b) Show that for a suitable choice ϵ, $y(t) = 0, \ \forall t > \epsilon$.

(c) Assuming $y(0^-) = 0$, what happens to $y(t)$ when, $\epsilon \to 0$ and $v_0 \to \frac{1}{\epsilon}$? Derive the mathematical expression applying this limit. Compare this solution to the input $v(t) = \delta(t)$.

(d) What will happen when $\epsilon \to 0$ and v_0 is constant?

112. Derive the differential equation governing the following two mechanical systems. The input to both these systems is the force $f(t)$, and the position $x(t)$ is the output; assume the initial conditions to $x(0^-), \dot{x}(0^-)$.



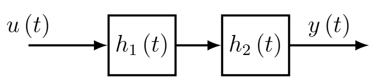
Find the expression for the step response of these two systems.

113. Consider a continuous-time LTI system with impulse response, $h(t) = e^{-2t}1(t)$. What is the output of this system to the following inputs using the convolution integral? (a) $e^{-2t}1(t)$; (b) e^{-2t}; (c) e^{-1t}; (d) e^{-4t}; and (e) $\cos(\omega t)$.

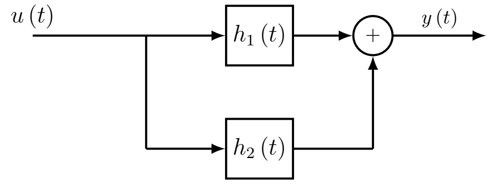
Now, obtain the expression for the output of the system for the above inputs using the system's transfer function $H(s)$.

114. Find the impulse response and transfer functions of the following composition of subsystems with individual impulse response $h_i(t)$.

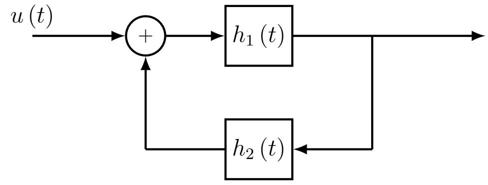
(a)



(b)



(c)



115. Consider the second order system, $\ddot{y}(t) + 2\zeta\omega_n\dot{y}(t) + \omega_n^2 y(t) = u(t)$. Find the impulse response of this system. Plot the impulse response of the system for $\omega_n = 1$ and the following values of the parameter ζ. (a) $\zeta = \sqrt{2}$; (b) $\zeta = 1$; (c) $\zeta = 0.5$; (d) $\zeta = 0$; (e) $\zeta = -0.5$; and (f) $\zeta = -1.0$; For each of these parameter values show the location of the poles of the corresponding transfer function.

References

[1] G. Strang *Introduction to linear algebra*. Wellesley-Cambridge Press Wellesley, MA, USA, 1993