Unit 3

LOS 1. Vector space

- can be made of different objects including functions
- consists of:
  - set of vectors i.e. column matrices or functions
  - set of scalars i.e. real numbers or complex numbers
- closed under vector addition and scalar multiplication i.e. if we have 2 vectors from the vector space, multiply them by a scalar and then add them, the final result is also a vector which is a member of the same vector space
- the zero vectors is a member of every vector space
- example of a vector space:
  - set of vectors: all 3×1 matrices
  - set of scalars: real numbers
  - if $u, v \in 3 \times 1$ matrices then $w = au + bv \in 3 \times 1$ matrix
- spaces:
  - null space
  - column space
  - row space
  - left null space

LOS 2. Linear independence

- The **set of vectors** $\{u_1, u_2, \ldots, u_n\}$ are linearly independent if $c_1u_1 + c_2u_2 + \ldots + c_nu_n = 0$ has the only solution $c_1 = c_2 = \ldots = c_n = 0$. In other words, no vector in the set can be written as a linear combination of the others.

LOS 3. Span, basis and dimension

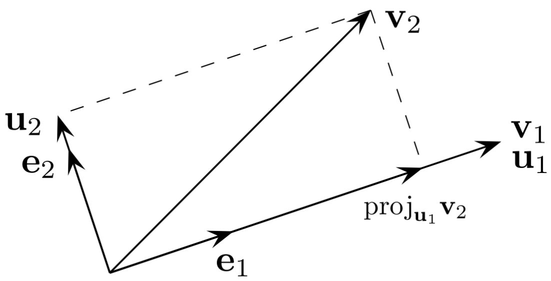
- **span**: $\{v_1, v_2, \ldots, v_n\}$ span a vector space consisting of all linear combinations of all linear combinations of $v_1, v_2, \ldots, v_n$. Example:
  - $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\}$
  - the above spans a vector space V of 3×1 matrices with 0 in the 3rd row
  - V is a vector subspace of all 3×1 matrices
- **basis**: a set of minimum number of vectors that span the space, possible basis for V (vector space of 3×1 matrices with 0 in the 3rd row)
  - $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ (orthonormal basis)

- $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\}$
- $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\}$

- **dimension**: dimension of a vector space is the least number of basis vectors required to form the vector space. In the above example, the dimension of the vector space V is 2.

LOS 3. Gram-Schmidt process:

- constructing an orthonormal basis for a vector space
- consider a basis $\{V_1, V_2, \ldots, V_n\}$ n linearly independent vectors that form a basis for an n-dimensional vector space, ordinary (not normalized nor orthogonal)
- Gram-Schmidt process provides a algorithm to convert an ordinary basis (not normalized, not orthogonal) to an orthonormal one $\{U_1, U_2, \ldots, U_n\}$
- steps:
  - find an orthonormal basis: U_2 is equal to the the original vector V_2 minus the projection of V_2 to $V_1 = U_1$. Here, we set U_1 as the reference vector and therefore equal to V_1



$$
\begin{aligned}
U_1 &= V_1 \\
U_2 &= V_2 - \frac{(U_1^T V_2)U_1}{U_1^T U_1} \\
U_3 &= V_3 - \frac{(U_1^T V_3)U_1}{U_1^T U_1} - \frac{(U_2^T V_3)U_2}{U_2^T U_2} \\
&\vdots \\
U_n &= \ \ldots
\end{aligned}
$$

  - normalized

$$
\begin{aligned}
U_1^* &= \frac{U_1}{(U_1^T U_1)^{\frac{1}{2}}} \\
&\vdots \\
U_n^* &= \frac{U_n}{(U_n^T U_n)^{\frac{1}{2}}}
\end{aligned}
$$

  - when calculating the new basis, we can ignore the magnitude as they will be normalized later, we can focus on forming vectors that have a 'nice' form i.e. whole numbers

- the span of a subset of the new vectors is the same as the span of the corresponding original vectors:

  $\mathrm{span}\{V_1, V_2, \ldots, V_k\} = \mathrm{span}\{U_1, U_2, \ldots, U_k\}$ where $k < n$

LOS 4. Null space:

- $\text{Null}(A)$ is a vector space of all column vectors x such that $Ax = 0$. Here, $\text{Null}(A)$ is a subspace of all 5 $\times$ 1 matrices
- $\text{Null}(A)$ is a vector space since the sum of 2 x's where $Ax = 0$ will also satisfy $Ax = 0$
- Method of finding $\text{Null}(A)$ the through example:
  - $A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}$
  - step 1: find the RREF of A

    $$\text{rref}(A) = \begin{pmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
  - step2: solve $\text{rref}(A)x = 0$

    $$\begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 = 0 \qquad & \rightarrow x_1 = 2x_2 + x_4 - 3x_5 \\ x_3 - 2x_4 - 2x_5 = 0 \qquad & \rightarrow x_3 = -2x_4 + 2x_5 \end{aligned}$$
  - parameterize x and split into vectors:

    $$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$
  - the split vectors form the basis of $\text{Null}(A)$
-
- observations:
  - x_1 and x_3 are in the pivot columns and not included in final vector equation, x_1 and x_3 are basic variables
  - x_2, x_4, and x_5 are free variables
  - $\dim(\text{Null}(A))$ is equal to the number of non-pivot columns or number of free variables

LOS 5. Application of the null space

- used to determine general solution of an underdetermined system of equations: $Ax = b$ with fewer equations than unknowns $\rightarrow$ the number of rows $<$ number of columns
- step:
  - let u be a general vector in $\text{Null}(A)$, any linear combination of the basis vectors of $\text{Null}(A)$
  - let v be any vector that solves $Ax = b$, the simplest approach would be to set the free variables to 0
  - $x = v + cu$ is general solution to $Ax = b$ where c is a scalar constant that can take any value
  - Proof: $Ax = A(u + v) = Au + Av = 0 + b = b$

LOS 6. Column space

- vector space spanned by the columns of the matrix
- form:

  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} = x_1 \begin{pmatrix} a \\ c \end{pmatrix} + x_2 \begin{pmatrix} b \\ d \end{pmatrix}$
- b has to be in the column space of the matrix A (b in $Ax = b$)
- find the basis for $\text{Col}(A)$ and its dimension
- steps:
  - find $\text{rref}(A)$
  - only the pivot columns are linearly independent, the other columns are linear combination of the pivot columns
  - the original pivot columns of A (before RREF) will be the basis for the column space
  - $\dim(\text{Col}(A))$ is equal to the number of pivot columns

LOS 7. Row space, left null space and rank

- subspaces for $A : m \times n$
  - $\text{Null}(A)$: subspace of all $n \times 1$ matrices
  - $\text{Col}(A)$: subspace of all $m \times 1$ matrices
  - $\text{Row}(A) = \text{Col}(A^T)$: subspace of all $n \times 1$ matrices
  - $\text{Leftnull}(A) = \text{Null}(A^T)$: subspace of all $m \times 1$ matrices
- observations:
  - $\dim(\text{Null}(A))$ = the number of non-pivot columns
  - $\dim(\text{Row}(A))$ = the number of pivot columns
  - $\dim(\text{Null}(A)) + \dim(\text{Row}(A)) = n$ (total number of columns of A)
  - vectors in the row space are orthogonal complement to the vector in the null space, they are orthogonal and when combined they form the entire vector space of $n \times 1$ matrices
  - $\dim(\text{Col}(A)) = \dim(\text{Row}(A))$ = the number of pivot columns = $\text{rank}(A)$: the number of linearly independent columns that the matrix has
  - the number of linearly independent rows and columns is always the same
- full rank: rank equal to the total number of columns

LOS 8. Orthogonal projections

- projecting a vector a big vector space unto a subspace of the original vector space
- consider:
  - V: n-dimensional vector space
  - W: p-dimensional subspace of V
  - $\{s_1, s_2, \ldots, s_p\}$: orthonormal basis for W
- steps:
  - let v be a vector in V
  - the orthogonal projection of v unto W:

    $v_{\text{proj}_W} = (v^T s_1)s_1 + (v^T s_2)s_2 + \ldots + (v^T s_p)s_p$

- intuition: we are projecting v along the basis of W
  - basis for V: $\{s_1, s_2, \ldots, s_p, t_1, t_2, \ldots, t_{n-p}\}$
  - if we have v in V, then: $v = a_1 s_1 + a_2 s_2 + \ldots + a_p s_p + b_1 t_1 + b_2 t_2 + \ldots + b_{n-p} t_{n-p}$ where a_i and b_i are scalar
  - then, the projection of v: $v_{\text{proj}_W} = a_1 s_1 + a_2 s_2 + \ldots + a_p s_p$ which is a piece of v in the subspace
- v_{proj_W} is the vector in W that is closest to v

LOS 9. Least-squares problem

- Consider:
  - data: $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$
  - x: exact
  - y: noisy data
  - $y = \beta_0 + \beta_1 x$, therefore $y_i = \beta_0 + \beta_1 x_i \; \forall i = 1, 2, \ldots, n$
- The matrix form: synonymous to $Ax = b$

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

- The problem above is overdetermined, we have more rows than columns, this means b is not in the column space of A. The best solution is therefore to project b into the column space of A.
- New problem: $Ax = b_{\text{proj}_{\text{Col}(A)}}$ where $b_{\text{proj}_{\text{Col}(A)}}$
  - consider $b = b_{\text{proj}_{\text{Col}(A)}} + (b - b_{\text{proj}_{\text{Col}(A)}})$
  - $(b - b_{\text{proj}_{\text{Col}(A)}})$ component is whatever's left after the projection, this is orthogonal to $\text{Col}(A)$(refer to the Gram-Schmidt section on projection of U unto V)
  - if $(b - b_{\text{proj}_{\text{Col}(A)}})$ is orthogonal to $\text{Col}(A)$, then it is also orthogonal to $\text{Row}(A^T)$ and therefore in null space of A^T, $\text{Null}(A^T)$ meaning $A^T(b - b_{\text{proj}_{\text{Col}(A)}}) = 0$
  - multiply both sides with A^T to get the normal equation:

    $A^T A x = A^T b$

    $x = (A^T A)^{-1} A^T b$
  - multiply both sides with A:

    $Ax = A(A^T A)^{-1} A^T b = b_{\text{proj}_{\text{Col}(A)}}$
  - $A(A^T A)^{-1} A^T$ is called the projection matrix which projects b into the column space of A