MATH1030 Orthogonal complement.

1. Recall the definition for the notion of *orthogonality* from the handout *Inner product, norm, and orthogonality*:

   Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *We say* $\mathbf{u}$ *is orthogonal to* $\mathbf{v}$, *and write* $\mathbf{u} \perp \mathbf{v}$, *if and only if* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

   Also reacll these basic properties of orthogonality:

   (a) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\mathbf{u} \perp \mathbf{v}$ *if and only if* $\mathbf{v} \perp \mathbf{u}$.
   (b) *Suppose* $\mathbf{u} \in \mathbb{R}^n$. *Then* $\mathbf{u} \perp \mathbf{u}$ *if and only if* $\mathbf{u} = \mathbf{0}_n$.
   (c) *Suppose* $\mathbf{u} \in \mathbb{R}^n$. *Then* ($\mathbf{u} \perp \mathbf{v}$ *for any* $\mathbf{v} \in \mathbb{R}^n$) *if and only if* $\mathbf{u} = \mathbf{0}_n$.
   (d) *Suppose* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then* $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ *if and only if* $\mathbf{u} \perp \mathbf{v}$.

2. **Definition. (Perp of of a subset of $\mathbb{R}^n$.)**

   Let S *be a set of vectors in* $\mathbb{R}^n$.

   The perp of S, *denoted by* $S^\perp$, *is defined to be* $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \perp \mathbf{u}$ *for any* $\mathbf{u} \in S\}$.

   Remarks.

   (a) In plain words, $S^\perp$ is the collection of those and only those vectors in $\mathbb{R}^n$ which are orthogonal to every vector in S.
   (b) i. $(\mathbb{R}^n)^\perp$ is the zero subspace of $\mathbb{R}^n$.
       ii. $(\{\mathbf{0}_n\})^\perp$ is $\mathbb{R}^n$ itself.
       iii. By logic, $\emptyset^\perp$ is $\mathbb{R}^n$ itself.

3. **Theorem (1).**

   Suppose S *is a set of vectors in* $\mathbb{R}^n$. *Then* $S^\perp$ *is a subspace of* $\mathbb{R}^n$.

4. **Proof of Theorem (1).**

   Suppose S is a set of vectors in $\mathbb{R}^n$.

   - [We want to check '$\mathbf{0}_n \in S^\perp$'. This amounts to checking 'for any $\mathbf{u} \in S$, $\mathbf{0}_n \perp \mathbf{u}$'.]
     For any $\mathbf{u} \in S$, we have $\langle \mathbf{0}_n, \mathbf{u} \rangle = 0$. Then $\mathbf{0}_n \perp \mathbf{u}$.
     It follows that $\mathbf{0}_n \in S^\perp$.
   - [We want to check 'for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, if $\mathbf{x} \in S^\perp$ and $\mathbf{y} \in S^\perp$ then $\mathbf{x} + \mathbf{y} \in S^\perp$'.]
     Pick any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Suppose $\mathbf{x} \in S$ and $\mathbf{y} \in S$.
     [We try to deduce '$\mathbf{x} + \mathbf{y} \in S^\perp$'.]
     We verify that for any $\mathbf{u} \in S$, $\mathbf{x} + \mathbf{y} \perp \mathbf{u}$':
       Pick any $\mathbf{u} \in S$.
       Since $\mathbf{x} \in S^\perp$, we have $\mathbf{x} \perp \mathbf{u}$. Then $\langle \mathbf{x}, \mathbf{u} \rangle = 0$.
       Similarly, since, $\mathbf{y} \in S^\perp$, we have $\langle \mathbf{y}, \mathbf{u} \rangle = 0$.
       Then $\langle \mathbf{x} + \mathbf{y}, \mathbf{u} \rangle = \langle \mathbf{x}, \mathbf{u} \rangle + \langle \mathbf{y}, \mathbf{u} \rangle = 0$.
       Therefore $\mathbf{x} + \mathbf{y} \perp \mathbf{u}$.
     It follows that $\mathbf{x} + \mathbf{y} \in S^\perp$.
   - As an exercise, verify that 'for any $\mathbf{x} \in \mathbb{R}^n$, for any $\alpha \in \mathbb{R}$, if $\mathbf{x} \in S^\perp$ then $\alpha\mathbf{x} \in S^\perp$'.

5. **Theorem (2).**

   Suppose S *is a set of vectors in* $\mathbb{R}^n$. *Then* S *is a subset of* $(S^\perp)^\perp$.

6. **Proof of Theorem (2).**

   Suppose S is a set of vectors in $\mathbb{R}^n$.

   [We want to verify that for any $\mathbf{y} \in \mathbb{R}^n$, if $\mathbf{y} \in S$ then $\mathbf{y} \in (S^\perp)^\perp$.]

   Pick any $\mathbf{y} \in \mathbb{R}^n$. Suppose $\mathbf{y} \in S$.

   [Reminder: We want to deduce that $\mathbf{y} \in (S^\perp)^\perp$.

   We now remind ourselves: according to definition, $(S^\perp)^\perp = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{z} \perp \mathbf{v}$ for any $\mathbf{v} \in S^\perp\}$.

So to verify '$\mathbf{y} \in (S^\perp)^\perp$' is the same as to verify '$\mathbf{y} \perp \mathbf{v}$ for any $\mathbf{v} \in S^\perp$'.

So we now proceed to verify that $\mathbf{y} \perp \mathbf{v}$ for any $\mathbf{v} \in S^\perp$.]

Pick any $\mathbf{v} \in S^\perp$. By assumption, $\mathbf{y} \in S$. Then $\mathbf{v} \perp \mathbf{y}$ by the definition of $S^\perp$. Therefore $\mathbf{y} \perp \mathbf{v}$.

(We have verified that $\mathbf{y} \perp \mathbf{v}$ for any $\mathbf{v} \in S^\perp$.)

Hence $\mathbf{y} \in (S^\perp)^\perp$.

It follows that S is a subset of $(S^\perp)^\perp$.

7. **Theorem (3).**

   Suppose A is an $(m \times n)$-matrix. Then the statements below hold:

   (a) $\mathcal{N}(A) = (\mathcal{R}(A))^\perp$. (b) $\mathcal{N}(A^t) = (\mathcal{C}(A))^\perp$.

8. **Proof of Theorem (3).**

   Suppose A is an $(m \times n)$-matrix.

   (a)
   - [We verify that 'for any $\mathbf{x} \in \mathbb{R}^n$, if $\mathbf{x} \in \mathcal{N}(A)$' then $\mathbf{x} \in (\mathcal{R}(A))^\perp$.]
     Pick any $\mathbf{x} \in \mathbb{R}^n$. Suppose $\mathbf{x} \in \mathcal{N}(A)$. [We want to deduce '$\mathbf{x} \in (\mathcal{R}(A))^\perp$'. This amounts to checking 'for any $\mathbf{u} \in (\mathcal{R}(A))$, $\mathbf{x} \perp \mathbf{u}$.']
     We verify that for any $\mathbf{u} \in (\mathcal{R}(A))$, $\mathbf{x} \perp \mathbf{u}$:
       Pick any $\mathbf{u} \in \mathcal{R}(A)$. [Ask: is it true that $\mathbf{x} \perp \mathbf{u}$?]
       Since $\mathbf{u} \in \mathcal{R}(A)$ and $\mathcal{R}(A) = \mathcal{C}(A^t)$, there exists some $\mathbf{s} \in R^m$ such that $\mathbf{u} = A^t\mathbf{s}$.
       Since $\mathbf{x} \in \mathcal{N}(A)$, we have $A\mathbf{x} = \mathbf{0}_m$.
       Now we have $\langle \mathbf{x}, \mathbf{u} \rangle = \mathbf{x}^t\mathbf{u} = \mathbf{x}^t A^t \mathbf{s} = (A\mathbf{x})^t\mathbf{s} = {\mathbf{0}_m}^t\mathbf{x} = 0$.
       Then $\mathbf{x} \perp \mathbf{u}$.
     Hence $\mathbf{x} \in (\mathcal{R}(A))^\perp$.
   - [We verify that 'for any $\mathbf{x} \in \mathbb{R}^n$, if $\mathbf{x} \in (\mathcal{R}(A))^\perp$ then $\mathbf{x} \in \mathcal{N}(A)$'.]
     Pick any $\mathbf{x} \in \mathbb{R}^n$. Suppose $\mathbf{x} \in (\mathcal{R}(A))^\perp$. [We want to deduce '$\mathbf{x} \in \mathcal{N}(A)$'. This amounts to checking '$A\mathbf{x} = \mathbf{0}_m$'.]
     We verify that $A\mathbf{x} = \mathbf{0}_m$:
       Since $\mathbf{x} \in (\mathcal{R}(A))^\perp$, we have $\mathbf{x} \perp \mathbf{u}$ for any $\mathbf{u} \in \mathcal{R}(A)$.
       Denote the columns of A^t by $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$. By definition, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m \in \mathcal{R}(A)$.
       Then ${\mathbf{u}_j}^t\mathbf{x} = \langle \mathbf{u}_j, \mathbf{x} \rangle = 0$ for each $j = 1, 2, \cdots, m$.
       We have $A\mathbf{x} = \begin{bmatrix} {\mathbf{u}_1}^t\mathbf{x} \\ \hline {\mathbf{u}_2}^t\mathbf{x} \\ \hline \vdots \\ \hline {\mathbf{u}_m}^t\mathbf{x} \end{bmatrix} = \mathbf{0}_m$.
     It follows that $\mathbf{x} \in \mathcal{N}(A)$.

   It follows that $\mathcal{N}(A) = (\mathcal{R}(A))^\perp$.

   (b) We have $\mathcal{R}(A^t) = \mathcal{C}(A)$ by definition. Then $\mathcal{N}(A^t) = (\mathcal{R}(A^t))^\perp = (\mathcal{C}(A))^\perp$.

9. **Theorem (4).**

   Let V be a subspace of $\mathbb{R}^n$. The statements below hold:

   (a) $V \cap V^\perp = \{\mathbf{0}_n\}$. *(The only vector which belongs to both V and $V^\perp$ is the zero vector in $\mathbb{R}^n$.)*

   (b) $\dim(V) + \dim(V^\perp) = \mathbb{R}^n$.

10. **Proof of Theorem (4).**

    Let V be a subspace of $\mathbb{R}^n$.

    (a) Pick any $\mathbf{x} \in \mathbb{R}^n$. Suppose $\mathbf{x} \in V \cap V^\perp$. Then $\mathbf{x} \in V$ and $\mathbf{x} \in V^\perp$.
    Since $\mathbf{x} \in V^\perp$, we have $\mathbf{x} \perp \mathbf{u}$ for any $\mathbf{u} \in V$. In particular, $\mathbf{x} \perp \mathbf{x}$ (because $\mathbf{x} \in V$).
    Then $\mathbf{x} = \mathbf{0}$.

(b) When V is the zero subspace of $\mathbb{R}^n$, $V^\perp$ is $\mathbb{R}^n$ and $\dim(V)+\dim(V^\perp)=n$. From now on we suppose V is not the zero subspace of $\mathbb{R}^n$.

Write $\dim(V)=p$. We have $p\geq 1$.

There is some basis for V, with p vectors in $\mathbb{R}^n$, say, $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p$.

Define the $(n\times p)$-matrix U by $U=\left[\ \mathbf{u}_1\ \middle|\ \mathbf{u}_2\ \middle|\ \cdots\ \middle|\ \mathbf{u}_p\ \right]$.

We have $V=\mathcal{C}(U)$.

By Theorem (3), we have $\mathcal{N}(U^t)=(\mathcal{C}(U))^\perp=V^\perp$.

Note that U^t is a $(p\times n)$-matrix.

By the Rank-Nullity Formula, we have $\dim(\mathcal{N}(U^t))+\dim(\mathcal{C}(U^t))=n$.

By definition, $\mathcal{R}(U)=\mathcal{C}(U^t)$.

Also recall that $\dim(\mathcal{C}(U))=\dim(\mathcal{R}(U))$, which the number of the leading ones in the reduced row-echelon form which is row-equivalent to U.

Then $\dim(V^\perp)=\dim(\mathcal{N}(U^t))=n-\dim(\mathcal{C}(U^t))=n-\dim(\mathcal{R}(U))=n-\dim(\mathcal{C}(U))=n-\dim(V)$.

Therefore $\dim(V)+\dim(V^\perp)=n$.

11. Recall the result (†) from the handout *Inequalities on dimension*:

(†) *Let W be a subspace of $\mathbb{R}^m$. Suppose $\mathbf{t}_1,\mathbf{t}_2,\cdots,\mathbf{t}_k$ be vectors in W.*
Then the statements below are logically equivalent:

($\sharp$) *$\mathbf{t}_1,\mathbf{t}_2,\cdots,\mathbf{t}_k$ constitute a basis for W.*

($\natural$) *$\dim(W)=k$, and $\mathbf{t}_1,\mathbf{t}_2,\cdots,\mathbf{t}_k$ are linearly independent.*

12. **Theorem (5).**

Let V be a subspace of $\mathbb{R}^n$, and $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p,\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ be vectors in $\mathbb{R}^n$.

Suppose $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p$ constitute a basis for V, and $\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ constitute a basis for $V^\perp$.

Then $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p,\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ constitute a basis for $\mathbb{R}^n$.

13. **Proof of Theorem (5).**

Let V be a subspace of $\mathbb{R}^n$, and $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p,\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ be vectors in $\mathbb{R}^n$.

Suppose $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p$ constitute a basis for V, and $\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ constitute a basis for $V^\perp$.

- [We verify that $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p,\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ are linearly independent.]
  Pick any $\alpha_1,\alpha_2,\cdots,\alpha_p,\beta_1,\beta_2,\cdots,\beta_{n-p}\in\mathbb{R}$.
  Suppose $\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\cdots+\alpha_p\mathbf{u}_p+\beta_1\mathbf{w}_1+\beta_2\mathbf{w}_2+\cdots+\beta_{n-p}\mathbf{w}_{n-p}=\mathbf{0}$.
  Write $\mathbf{u}=\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\cdots+\alpha_p\mathbf{u}_p$, $\mathbf{w}=\beta_1\mathbf{w}_1+\beta_2\mathbf{w}_2+\cdots+\beta_{n-p}\mathbf{w}_{n-p}$. Then $\mathbf{u}+\mathbf{w}=\mathbf{0}_n$.
  By definition, $\mathbf{u}\in V$. Then $\mathbf{w}=-\mathbf{u}\in V$.
  By definition, $\mathbf{w}\in V^\perp$. Then $\mathbf{u}=-\mathbf{w}\in V^\perp$.
  Therefore $\mathbf{u}\in V$ and $\mathbf{u}\in V^\perp$. Hence $\mathbf{u}\in V\cap V^\perp$.
  Then $\alpha_1\mathbf{u}_1+\alpha_2\mathbf{u}_2+\cdots+\alpha_p\mathbf{u}_p=\mathbf{u}=\mathbf{0}_n$.
  Recall that $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p$ are linearly independent. Then $\alpha_1=\alpha_2=\cdots=\alpha_p=0$.
  Similarly, $\mathbf{w}\in V$ and $\mathbf{w}\in V^\perp$. Then $\mathbf{w}\in V\cap V^\perp$. Therefore $\beta_1\mathbf{w}_1+\beta_2\mathbf{w}_2+\cdots+\beta_{n-p}\mathbf{w}_{n-p}=\mathbf{w}=\mathbf{0}_n$.
  Recall that $\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ are linearly independent. Then $\beta_1=\beta_2=\cdots=\beta_{n-p}=0$.
- We have now proved that $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p,\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ are n linearly independent vectors in $\mathbb{R}^n$, which is an n-dimensional subspace of $\mathbb{R}^n$.
  It follows from the result (†) that $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_p,\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_{n-p}$ constitute a basis for $\mathbb{R}^n$.

14. **Definition. (Orthogonal complement of a subspace of $\mathbb{R}^n$.)**

Suppose V is a subspace of $\mathbb{R}^n$. Then $V^\perp$ is called the orthogonal complement of V in $\mathbb{R}^n$.

15. **Theorem (6).**

Suppose V is a subspace of $\mathbb{R}^n$. Then for any $\mathbf{x}\in\mathbb{R}^n$, there exist some unique $\mathbf{u}\in V$, $\mathbf{w}\in V^\perp$ such that $\mathbf{v}=\mathbf{u}+\mathbf{w}$.

Remarks.

(a) What the conclusion in this statement is saying is that both statements below are true:

(E) *For any* $\mathbf{x} \in \mathbb{R}^n$, *there exist some unique* $\mathbf{u} \in V$, $\mathbf{w} \in V^\perp$ *such that* $\mathbf{v} = \mathbf{u} + \mathbf{w}$.

(U) *For any* $\mathbf{x} \in \mathbb{R}^n$, *for any* $\mathbf{u}, \mathbf{u}' \in V$, $\mathbf{w}, \mathbf{w}' \in V^\perp$ *if* $\mathbf{v} = \mathbf{u} + \mathbf{w}$ *and* $\mathbf{v} = \mathbf{u}' + \mathbf{w}'$ *then* $\mathbf{u} = \mathbf{u}'$ *and* $\mathbf{w} = \mathbf{w}'$.

(b) The statement (E) is called the 'existence part' (of the conclusion) in Theorem (6). In plain words, it says that every vector in $\mathbb{R}^n$ 'admits' at least one 'decomposition' as a sum of two of vectors, one from V and the other from $V^\perp$.

(c) The statement (U) is called the 'uniqueness part' (of the conclusion) in Theorem (6). In plain words, it says that every vector in $\mathbb{R}^n$ 'admits' at most one 'decomposition' as a sum of two of vectors, one from V and the other from $V^\perp$.

16. **Proof of Theorem (6).**

Suppose V is a subspace of $\mathbb{R}^n$. Write $\dim(V) = p$.

Note that $V^\perp$ is a subspace of $\mathbb{R}^n$, and $\dim(V^\perp) = n - p$.

Pick any $\mathbf{x} \in \mathbb{R}^n$.

- Pick some basis for V, say, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$.
  Pick some basis for $V^\perp$, say, $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{n-p}$.
  By Theorem (5), $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p, \mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{n-p}$ constitute a basis for $\mathbb{R}^n$.
  There exist some $\alpha_1, \alpha_2, \cdots, \alpha_p, \beta_1, \beta_2, \cdots, \beta_{n-p} \in \mathbb{R}$ such that

$$\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p + \beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \cdots + \beta_{n-p}\mathbf{w}_{n-p}.$$

  Define $\mathbf{u} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_p\mathbf{u}_p$, $\mathbf{w} = \beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \cdots + \beta_{n-p}\mathbf{w}_{n-p}$.
  By definition, $\mathbf{u} \in V$, $\mathbf{w} \in V^\perp$, and $\mathbf{v} = \mathbf{u} + \mathbf{w}$.
- Pick any $\mathbf{u}, \mathbf{u}' \in V$, $\mathbf{w}, \mathbf{w}' \in V^\perp$. Suppose $\mathbf{v} = \mathbf{u} + \mathbf{w}$ and $\mathbf{v} = \mathbf{u}' + \mathbf{w}'$.
  Then $\mathbf{u} + \mathbf{w} = \mathbf{v} = \mathbf{u}' + \mathbf{w}'$.
  Therefore $\mathbf{u} - \mathbf{u}' = \mathbf{w}' - \mathbf{w}$.
  Since $\mathbf{u}, \mathbf{u}' \in V$, we have $\mathbf{u} - \mathbf{u}' \in V$.
  Since $\mathbf{w}, \mathbf{w}' \in V^\perp$, we have $\mathbf{w}' - \mathbf{w} \in V^\perp$.
  We have $\mathbf{u} - \mathbf{u}' \in V$ and $\mathbf{u} - \mathbf{u}' \in V^\perp$. Therefore $\mathbf{u} - \mathbf{u}' \in V \cap V^\perp$. By Theorem (4), we have $\mathbf{u} - \mathbf{u}' = \mathbf{0}_n$.
  Hence $\mathbf{u} = \mathbf{u}'$.
  Now we have $\mathbf{w}' - \mathbf{w} = \mathbf{u} - \mathbf{u}' = \mathbf{0}_n$ also. Then $\mathbf{w} = \mathbf{w}'$.

17. Recall the result (‡) from the handout *Inequalities on dimension*:

(‡) *Let* W_1, W_2 *be subspaces of* $\mathbb{R}^m$. *Suppose* W_1 *is a subspace of* W_2.
Then $\dim(W_1) \leq \dim(W_2)$. *Equality holds if and only if* $W_1 = W_2$.

18. **Theorem (7).**

Suppose V *is a subspace of* $\mathbb{R}^n$. *Then* $(V^\perp)^\perp = V$.

Proof of Theorem (7).

Suppose V is a subspace of $\mathbb{R}^n$.

By Theorem (2), V is a subspace of $(V^\perp)^\perp$.

By Theorem (4), $\dim(V) + \dim(V^\perp) = n$. Also, $\dim(V^\perp) + \dim((V^\perp)^\perp) = n$.

Then $\dim(V) = n - \dim(V^\perp) = \dim((V^\perp)^\perp)$.

It follows that $V = (V^\perp)^\perp$.

19. With the help of Theorem (7), we may extend Theorem (3) into the result below:

Theorem (8).

Suppose A *is an* $(m \times n)$*-matrix. Then the statements below hold:*

(a) $\mathcal{N}(A) = (\mathcal{R}(A))^\perp$. (b) $\mathcal{N}(A^t) = (\mathcal{C}(A))^\perp$. (c) $(\mathcal{N}(A))^\perp = \mathcal{R}(A)$. (d) $(\mathcal{N}(A^t))^\perp = \mathcal{C}(A)$.

Proof of Theorem (8).

Suppose A is an $(m \times n)$-matrix.

By Theorem (3), we have $\mathcal{N}(A) = (\mathcal{R}(A))^\perp$ and $\mathcal{N}(A^t) = (\mathcal{C}(A))^\perp$.

By Theorem (7), we have $(\mathcal{N}(A))^\perp = ((\mathcal{R}(A))^\perp)^\perp = \mathcal{R}(A)$.

We also have $(\mathcal{N}(A^t))^\perp = ((\mathcal{C}(A))^\perp)^\perp = \mathcal{C}(A)$.

Remark. The set equality $(\mathcal{N}(A^t))^\perp = \mathcal{C}(A)$ is logically equivalent to the statement below:

For any $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{b} \in \mathcal{C}(A)$ if and only if $\mathbf{b} \in (\mathcal{N}(A^t))^\perp$.

This set equality is usually re-formulated as a result known as Fredholm's Alternatives.

20. **Theorem (9). (Fredholhm's Alternatives.)**

Suppose A is an $(m \times n)$-matrix.

Then, for any $\mathbf{b} \in \mathbb{R}^m$, exactly one of the statements below is true:

(a) $\mathbf{b} \in \mathcal{C}(A)$. (b) $\mathbf{b} \notin (\mathcal{N}(A^t))^\perp$.

Remarks.

- '$\mathbf{b} \in \mathcal{C}(A)$' is logically equivalent to the statement '$\mathcal{LS}(A, \mathbf{b})$ is consistent'.
- '$\mathbf{b} \in (\mathcal{N}(A^t))^\perp$' is logically equivalent to the statement 'every vector in the null space of A^t is orthogonal to $\mathbf{b}$'.

  Recall that the null space of A^t is the solution set of $\mathcal{LS}(A^t, \mathbf{0}_n)$.

  So '$\mathbf{b} \in (\mathcal{N}(A^t))^\perp$' is logically equivalent to the statement 'every solution of the system $\mathcal{LS}(A^t, P\mathbf{0}_m)$ is orthogonal to $\mathbf{b}$'.

This allows the re-formulation of Fredholhm's Alternatives in terms of systems of linear equations.

21. **Corollary to Theorem (9). (Fredholm's Alternatives, in terms of systems of linear equations.)**

Suppose A is an $(m \times n)$-matrix, and $\mathbf{b}$ is a vector in $\mathbb{R}^m$.

Then exactly one of the statements below is true:

(a) *$\mathcal{LS}(A, \mathbf{b})$ is consistent.*

(b) *Some non-trivial solution of the homogeneous system $\mathcal{LS}(A^t, \mathbf{0}_n)$ is not orthogonal to $\mathbf{b}$.*

22. Recall the result below from the handout *Duality between spanning and linear independence*:

Let A be an $(m \times n)$-matrix. Suppose K is a non-singular $(m \times m)$-square matrix. Then the equalities below hold:

(a) $\mathcal{N}(A) = \mathcal{N}(KA)$. (b) $\mathcal{R}(A) = \mathcal{R}(KA)$.

This result can be re-formulated as:

($\star$) *Let A, B be $(m \times n)$-matrices. Suppose A is row-equivalent to B. Then $\mathcal{N}(A) = \mathcal{N}(B)$ and $\mathcal{R}(A) = \mathcal{R}(B)$.*

It turns out that the converse of this result is also true.

23. For the moment, we take for granted the validity of Lemma (10), which will be proved later.

Lemma (10).

Let P, Q be $(m \times n)$-square matrices. Suppose P, Q are reduced row-echelon forms, and $\mathcal{R}(P) = \mathcal{R}(Q)$. Then $P = Q$.

24. **Lemma (11).**

Let A, B be $(m \times n)$-matrices. Suppose $\mathcal{R}(A) = \mathcal{R}(B)$. Then A is row-equivalent to B.

Proof of Lemma (11).

Let A, B be $(m \times n)$-matrices. Suppose $\mathcal{R}(A) = \mathcal{R}(B)$.

Denote by A' the reduced row-echelon form which is row-equivalent to A.

Denote by B' the reduced row-echelon form which is row-equivalent to B.

We have $\mathcal{R}(A) = \mathcal{R}(A')$ and $\mathcal{R}(B) = \mathcal{R}(B')$.

Now, by assumption, we have $\mathcal{R}(A') = \mathcal{R}(A) = \mathcal{R}(B) = \mathcal{R}(B')$.

Then, by Lemma (10), we have $A' = B'$. It follows that A is row-equivalent to B.

25. **Lemma (12).**

Let A, B be $(m \times n)$-matrices. Suppose $\mathcal{N}(A) = \mathcal{N}(B)$. Then A is row-equivalent to B.

Proof of Lemma (12).

Let A, B be $(m \times n)$-matrices. Suppose $\mathcal{N}(A) = \mathcal{N}(B)$.

Then $(\mathcal{R}(A))^\perp = \mathcal{N}(A) = \mathcal{N}(B) = (\mathcal{R}(B))^\perp$.

Therefore $\mathcal{R}(A) = ((\mathcal{R}(A))^\perp)^\perp = ((\mathcal{R}(B))^\perp)^\perp = \mathcal{R}(B)$.

26. We combine Lemma (11), Lemma (12) and the result $(\star)$ into the result below:

Theorem (13).

Let A, B be $(m \times n)$-matrices. The statements below are logically equivalent:

(a) *A is row-equivalent to B.* (b) $\mathcal{N}(A) = \mathcal{N}(B)$. (c) $\mathcal{R}(A) = \mathcal{R}(B)$.

27. **Corollary to Theorem (13).**

Let A, B be $(m \times n)$-matrices. The statements below are logically equivalent:

(a) *A is row-equivalent to B.*

(b) *$\mathcal{LS}(A, \mathbf{0})$ is equivalent to $\mathcal{LS}(B, \mathbf{0})$ (in the sense that their solution sets are equal to each other as sets).*

28. **Proof of Lemma (10).**

Let P, Q be $(m \times n)$-square matrices. Suppose P, Q are reduced row-echelon forms, and $\mathcal{R}(P) = \mathcal{R}(Q)$.

Write $k = \dim(\mathcal{R}(P))$.

(a) The respective numbers of non-zero rows of $\mathcal{R}(P)$ and of $\mathcal{R}(Q)$ are the same: it is k.
Suppose the pivot columns of P, from left to right, are the c_1-th, c_2-th, ..., c_k-th columns.
Suppose the pivot columns of Q, from left to right, are the d_1-th, d_2-th, ..., d_k-th columns.

(b) i. Denote the non-zero columns of P^t, from left to right, by $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_k$.
- By definition, for each $j = 1, 2, \cdots, k$, the c_j-th entry of $\mathbf{p}_j$ is 1. Whenever $i \neq j$, the c_i-th entry of $\mathbf{p}_j$ is 0. Whenever $\ell < c_j$, the ℓ-th entry of $\mathbf{p}_j$ is 0.

ii. Denote the non-zero columns of Q^t, from left to right, by $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k$.
- By definition, for each $j = 1, 2, \cdots, k$, the d_j-th entry of $\mathbf{q}_j$ is 1. Whenever $i \neq j$, the d_i-th entry of $\mathbf{q}_j$ is 0. Whenever $\ell < d_j$, the ℓ-th entry of $\mathbf{q}_j$ is 0.

(c) i. We have $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\}) = \mathcal{R}(C) = \mathcal{R}(D) = \mathsf{Span}\ (\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\})$. Then $c_1 = d_1$. (Why?)
Each of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$.
For each $j = 1, 2, \cdots, k$, there exist some $\alpha_{j1}, \alpha_{j2}, \cdots, \alpha_{jk} \in \mathbb{R}$ such that $\mathbf{u}_j = \alpha_{j1}\mathbf{v}_1 + \alpha_{j2}\mathbf{v}_2 + \cdots + \alpha_{jk}\mathbf{v}_k$.
We have $\alpha_{k1} = 0$; otherwise, we would have $d_1 < d_k \leq c_1$, which is impossible.
Similarly, we deduce $\alpha_{21} = \alpha_{31} = \cdots = \alpha_{k-1,1} = 0$.
Then $\mathbf{u}_2, \mathbf{u}_3, \cdots, \mathbf{u}_k \in \mathsf{Span}\ (\{\mathbf{v}_2, \mathbf{v}_3, \cdots, \mathbf{v}_k\})$.

ii. Repeating the above argument, we also deduce that $\mathbf{v}_2, \mathbf{v}_3, \cdots, \mathbf{v}_k \in \mathsf{Span}\ (\{\mathbf{u}_2, \mathbf{u}_3, \cdots, \mathbf{u}_k\})$.
Then $\mathsf{Span}\ (\{\mathbf{u}_2, \mathbf{u}_3, \cdots, \mathbf{u}_k\}) = \mathsf{Span}\ (\{\mathbf{v}_2, \mathbf{v}_3, \cdots, \mathbf{v}_k\})$. It follows that $c_2 = d_2$.

iii. Repeating this argument again and again, we deduce that for each $j = 1, 2, \cdots, k$, $c_j = d_j$, and
$\mathsf{Span}\ (\{\mathbf{u}_j, \mathbf{u}_{j+1}, \cdots, \mathbf{u}_k\}) = \mathsf{Span}\ (\{\mathbf{v}_j, \mathbf{v}_{j+1}, \cdots, \mathbf{v}_k\})$.

(d) i. Now we have $\mathsf{Span}\ (\{\mathbf{u}_k\}) = \mathsf{Span}\ (\{\mathbf{v}_k\})$. The first non-zero entries of $\mathbf{u}_k$, $\mathbf{v}_k$ are 1. Then $\mathbf{u}_k = \mathbf{v}_k$.

ii. We have $\mathbf{u}_{k-1} = \alpha_{k-1,k-1}\mathbf{v}_{k-1} + \alpha_{k-1,k}\mathbf{v}_k$. The first non-zero entries of $\mathbf{u}_{k-1}$, $\mathbf{v}_{k-1}$ are 1. Then $\alpha_{k-1,k-1} = 1$. The c_k-th entries of $\mathbf{u}_{k-1}$, $\mathbf{v}_{k-1}$, $\mathbf{v}_k$ are $0, 0, 1$ respectively. Then $\alpha_{k-1,k} = 0$. Therefore $\mathbf{u}_{k-1} = \mathbf{v}_{k-1}$.

iii. Repeating this argument, we deduce in succession that $\mathbf{u}_{k-2} = \mathbf{v}_{k-2}$, ..., $\mathbf{u}_2 = \mathbf{v}_2$, $\mathbf{u}_1 = \mathbf{v}_1$.

It follows that $P = Q$.