MATH1030 Existence and uniqueness of solutions for a system of linear equations whose coefficient matrix is a square matrix

1. **Lemma (θ). (Invertibility as a sufficient condition for existence and uniqueness of solution for systems of linear equations.)**

   Let A be an $(n \times n)$-square matrix.

   Suppose A is invertible.

   Then, for any vector $\mathbf{b}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{b})$ has one and only one solution, namely, '$\mathbf{x} = A^{-1}\mathbf{b}$'.

   Proof of Lemma (θ).

   Suppose A is invertible. Then its matrix inverse A^{-1} is well-defined as an $(n \times n)$-square matrix.

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

   - We have $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b}$.
     Hence '$\mathbf{x} = A^{-1}\mathbf{b}$' is a solution of the system $\mathcal{LS}(A, \mathbf{b})$.
   - Suppose '$\mathbf{x} = \mathbf{v}$' is a solution of the system $\mathcal{LS}(A, \mathbf{b})$.
     [Ask: Can it happen that $\mathbf{v}$ fails to be $A^{-1}\mathbf{b}$? (We hope not.)]
     Then $\mathbf{b} = A\mathbf{v}$. Therefore $A^{-1}\mathbf{b} = A^{-1}(A\mathbf{v}) = (A^{-1}A)\mathbf{v} = I_n\mathbf{v} = \mathbf{v}$.

   Hence '$\mathbf{x} = A^{-1}\mathbf{b}$' is the one and only one solution of the system $\mathcal{LS}(A, \mathbf{b})$.

2. **Question.** Is invertibility a necessary condition for existence and uniqueness of solution for systems of linear equations?

   What we are asking is whether the converse of Lemma (θ), as formulated below, is true:

   > '*Let A be an $(n \times n)$-square matrix.*
   > *Suppose that, for any vector $\mathbf{b}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{b})$ has one and only one solution.*
   > *Then A is invertible.*'

   Answer. The answer to this question is *yes*.

   Actually it takes 'much less' for a square matrix to be necessarily invertible. This is the point of Lemma (ι) and Lemma (κ).

3. **Lemma (ι).**

   Let A be an $(n \times n)$-square matrix.

   Suppose that, for any vector $\mathbf{c}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{c})$ has at most one solution.

   Then A is invertible.

   Lemma (κ).

   Let A be an $(n \times n)$-square matrix.

   Suppose that, for any vector $\mathbf{d}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{d})$ has at least one solution.

   Then A is invertible.

4. **Proof of Lemma (ι).**

   Let A be an $(n \times n)$-square matrix.

   Suppose that, for any vector $\mathbf{c}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{c})$ has at most one solution.

   Then, in particular, the homogeneous system $\mathcal{LS}(A, \mathbf{0})$ has at most one solution.

   Recall that $\mathcal{LS}(A, \mathbf{0})$ has at least one solution, namely, the trivial solution.

   Then $\mathcal{LS}(A, \mathbf{0})$ has exactly one solution, namely, the trivial solution. Therefore, by Lemma (1), A is non-singular.

   Hence, by Theorem (B), A is invertible.

5. **Proof of Lemma (κ).**

   Let A be an $(n \times n)$-square matrix.

   Suppose that, for any vector $\mathbf{d}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{d})$ has at least one solution.

   Denote by $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ the respective columns of I_n from left to right. (So $I_n = [\ \mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n\ ]$.)

By assumption, for each $j = 1, 2, \cdots, n$, the system $\mathcal{LS}(A, \mathbf{e}_j)$ has at least one solution, say, '$\mathbf{x} = \mathbf{g}_j$'. So by definition, $A\mathbf{g}_j = \mathbf{e}_j$.

Define $G = [\ \mathbf{g}_1 \mid \mathbf{g}_2 \mid \cdots \mid \mathbf{g}_n\ ]$

We verify that $AG = I_n$:

- We have $AG = A[\ \mathbf{g}_1 \mid \mathbf{g}_2 \mid \cdots \mid \mathbf{g}_n\ ] = [\ A\mathbf{g}_1 \mid A\mathbf{g}_2 \mid \cdots \mid A\mathbf{g}_n\ ] = [\ \mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n\ ] = I_n$.

Now by Theorem (A), G is non-singular. Therefore by Theorem (B), G is invertible. Furthermore, A is a matrix inverse of G. (Why?) Hence A is also invertible.

6. We can summarize our discussion above in Lemma (λ).

   Lemma (λ).

   Let A be an $(n \times n)$-square matrix.

   The statements below are logically equivalent:

   (a) *A is invertible.*

   (b) *For any vector $\mathbf{b}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{b})$ has one and only one solution, namely, '$\mathbf{x} = A^{-1}\mathbf{b}$'.*

   (c) *For any vector $\mathbf{c}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{c})$ has at least one solution.*

   (d) *For any vector $\mathbf{d}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{d})$ has at most one solution.*

   We combine Lemma (λ) and Theorem (C) to obtain an upgrade of Theorem (C) in the form of Theorem (E).

7. **Theorem (E). (Various re-formulations for the notions of non-singularity and invertibility.)**

   Let A be an $(n \times n)$-matrix. The statements below are logically equivalent:

   (a) *A is non-singular.*

   (b) *For any vector $\mathbf{v}$ in $\mathbb{R}^n$, if $A\mathbf{v} = \mathbf{0}$ then $\mathbf{v} = \mathbf{0}$.*

   (c) *The trivial solution is the only solution of the homogeneous system $\mathcal{LS}(A, \mathbf{0})$.*

   (d) *A is row-equivalent to I_n.*

   (e) *A is invertible.*

   (f) *There exists some $(n \times n)$-square matrix H such that $HA = I_n$.*

   (g) *There exists some $(n \times n)$-square matrix G such that $AG = I_n$.*

   (h) *For any vector $\mathbf{b}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{b})$ has one and only one solution, namely, '$\mathbf{x} = A^{-1}\mathbf{b}$'.*

   (i) *For any vector $\mathbf{c}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{c})$ has at least one solution.*

   (j) *For any vector $\mathbf{d}$ in $\mathbb{R}^n$, the system $\mathcal{LS}(A, \mathbf{d})$ has at most one solution.*

   Now suppose A is non-singular, with a sequence of row operations

   $$A = C_1 \xrightarrow[\rho_1]{} C_2 \xrightarrow[\rho_2]{} \cdots\cdots \xrightarrow[\rho_{p-2}]{} C_{p-1} \xrightarrow[\rho_{p-1}]{} C_p = I_n,$$

   and with H_k being the row-operation matrix corresponding to ρ_k for each k. Then $[I_n|A^{-1}]$ is the resultant of the application of the same sequence of row operations $\rho_1, \rho_2, \cdots, \rho_{p-1}$ starting from $[A|I_n]$:

   $$[A|I_n] = [C_1|I_n] \xrightarrow[\rho_1]{} [C_2|H_1] \xrightarrow[\rho_2]{} [C_3|H_2H_1] \xrightarrow[\rho_3]{} \cdots\cdots \xrightarrow[\rho_{p-2}]{} [C_{p-1}|H_{p-2}\cdots H_2H_1] \xrightarrow[\rho_{p-1}]{} [C_p|H_{p-1}\cdots H_2H_1] = [I_n|A^{-1}].$$

   Moreover, A^{-1} and A are respectively given as products of row-operation matrices by

   $$A^{-1} = H_{p-1}\cdots H_2H_1, \qquad A = {H_1}^{-1}{H_2}^{-1}\cdots{H_{p-1}}^{-1}.$$