MA1101R

Jia Cheng

September 2020

The textbook used is Linear Algebra, Concepts and Techniques on Euclidean Spaces 2nd Edition.

Uniqueness of row-reduced echelon forms   We provide a proof for the uniqueness of RREFs.

Let A be an augmented $m \times (n+1)$ matrix representing a linear system S in n variables. Let R_1, R_2 be 2 matrices in RREF equivalent to A. (Here equivalence is known as row equivalence.) By symmetry and transitivity of equivalence, we have $R_1 \equiv R_2$.

Claim: The pivot (and non-pivot) columns of R_1, R_2 are the same.

Suppose not. WLOG (if not we simply swap the roles of R_1 and R_2), there exists an index $i \in \{1, 2, \ldots, n\}$ such that the i-th column is a pivot column in R_2 but not a pivot in R_1. We then form a new $2m \times (n+1)$ matrix by stacking R_2 below R_1.
We note that the linear system S' represented by this stacked matrix is equivalent to the original linear system S because: Let S_1 be the linear system described by R_1, S_2 be the linear system described by R_2. Then $S' = S_1 \wedge S_2 = S \wedge S = S$.
By performing Gauss-Jordan elimination again, we see that the resulting matrix consists of strictly more pivot columns than R_1 originally had. In particular, all the pivot columns of R_1 are still pivot columns, plus column i and possibly some more columns. This is a contradiction since the general set of solutions had a strictly smaller dimension. Even if we ignore dimension, we can make the following argument. Let $C \subseteq \{1, 2, \ldots, n\}$ be the indices of non-pivot columns in R_1, C' be the number of non-pivot columns in the RREF of the stacked matrix. Then $C' \subsetneq C$.

For any fixed tuple corresponding to the non-pivot columns of C', there are infinitely many solutions to be found by S_1, since $C - C' \neq \emptyset$ and we can vary any variable corresponding to the non-pivot columns in $C - C'$. Whereas in system S', fixed the variables corresponding to C' forces a single tuple of solutions. This suggests that S_1 and S are not equivalent, which is a contradiction.

Hence R_1, R_2 have the same pivot columns. Now we consider the first row of R_1, R_2.

Claim: The first row of these 2 matrices must be the same.

We already know that cells in other pivot columns must be zero for both matrices, by definition of RREF. Now consider the cells corresponding to the non-pivot columns. We see that the variable corresponding to the pivot cell can be expressed in terms of the variables corresponding to the non-pivot cells. In other words, for both R_1, R_2, the coefficients w.r.t. those non-pivot variables must be the same. This says that all cells in the first row of R_1, R_2 must be the same.

By repeating this argument for every row, we prove that all rows of R_1, R_2 are equal and hence $R_1 = R_2$.

Matrix multiplication   Let $A \in \mathbb{R}^{m,p}, B \in \mathbb{R}^{p,n}$.

We represent the columns of A as $a_i, 1 \le i \le p$ and the rows of A as $a'_i, 1 \le i \le m$. Similarly, we represent the columns of B as b_i and rows as b'_i.

Then, $A = \begin{bmatrix} a_1 & a_2 & \dots & a_p \end{bmatrix} = \begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_m \end{bmatrix}, B = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix} = \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_p \end{bmatrix}$

The following are 2 ways to view multiplying AB as blocks.

$$AB = A \cdot \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \dots & Ab_n \end{bmatrix}$$

We can let $n = 1$ and observe that $AB = Ab_1$, hence the above is a simple generalization of this observation.

$$AB = \begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_m \end{bmatrix} \cdot B = \begin{bmatrix} a'_1 B \\ a'_2 B \\ \vdots \\ a'_m B \end{bmatrix}$$

Similarly, we can let $m = 1$ and see that $AB = a'_1 B$, and generalize from there.

Next, we let $n = 1$, such that b'_i are to be viewed as real numbers (strictly speaking they are vectors in $\mathbb{R}^1$, isomorphic to $\mathbb{R}$). Then,

$$A\mathbf{b} = \begin{bmatrix} a_1 & a_2 & \dots & a_p \end{bmatrix} \cdot \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_p \end{bmatrix} = \sum_{1 \le i \le p} b'_i \cdot \mathbf{a}_i$$

Similarly, if we let $m = 1$, we can multiply a row vector with a matrix as follows. Here a_i are real numbers.

$$\mathbf{a}B = \begin{bmatrix} a_1 & a_2 & \dots & a_p \end{bmatrix} \cdot \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_p \end{bmatrix} = \sum_{1 \le i \le p} a_i \cdot \mathbf{b}'_i$$

Transition matrices We provide another derivation of the change of coordinate/basis matrix.

Suppose $S = \{u_1, \dots, u_k\}, T = \{v_1, \dots, v_k\}$ are 2 bases of vsp V. Let $w \in V$. Let $[w]_S$ denote the coordinate vector of w w.r.t basis S, and define $[w]_T$ similarly. Then

$$w = \begin{bmatrix} u_1 & u_2 & \dots & u_k \end{bmatrix} [u]_S = \begin{bmatrix} v_1 & v_2 & \dots & v_k \end{bmatrix} [u]_T$$

which implies

$$[u]_T = \begin{bmatrix} v_1 & v_2 & \dots & v_k \end{bmatrix}^{-1} \begin{bmatrix} u_1 & u_2 & \dots & u_k \end{bmatrix} [u]_S$$

Letting $A := \begin{bmatrix} u_1 & u_2 & \dots & u_k \end{bmatrix} [u], B := \begin{bmatrix} v_1 & v_2 & \dots & v_k \end{bmatrix}$, we see that

$$[u]_T = B^{-1} \begin{bmatrix} u_1 & u_2 & \dots & u_k \end{bmatrix} [u]_S = \begin{bmatrix} B^{-1}u_1 & B^{-1}u_2 & \dots & B^{-1}u_k \end{bmatrix} [u]_S$$

Since for all vectors w', $w' = B[w']_T$, we see that $[w']_T = B^{-1}w'$, hence for each i, we have $B^{-1}u_i = [u_i]_T$.

Alternative proof of thm 4.2.8   Claim: Given $A \in \mathbb{R}^{m,n}, B \in \mathbb{R}^{n,p}, rank(AB) \leq \min\{rank(A), rank(B)\}$.

By considering column space, $rank(AB) = \dim\{ABx : x \in \mathbb{R}^{p,1}\} \leq \dim\{Ax : x \in \mathbb{R}^{n,1}\} = rank(A)$ since for each $x \in \mathbb{R}^{p,1}$, $Bx \in \mathbb{R}^{n,1}$.

Similarly, by considering row space, $rank(AB) = \dim\{xAB : x \in \mathbb{R}^{1,m}\} \leq \dim\{xB : x \in \mathbb{R}^{1,n}\} = rank(B)$

Proof of Gram-Schmidt Process   Claim: The Gram-Schmidt Process is correct and produces from a basis $\{u_i : 1 \leq i \leq n\}$ of vsp V an orthogonal basis $\{v_i : 1 \leq i \leq n\}$.

In the proof we will make use of this lemma. Given a set of vectors S and subset $A \subseteq S$. Suppose there is a set B for which $span(A) = span(B)$, then we can do the replacement $S - A \cup B$ such that $span(S - A \cup B) = span(S)$.

We do induction on i, where $P(i)$ states that $span\{v_1, \ldots, v_i\} = span\{u_1, \ldots, u_i\}$. Since $v_1 = u_1$, the base case is true.

Suppose $P(k)$ is true. We note that the $k + 1$-th vector, $v_{k+1} = u_{k+1} - \sum_{1 \leq i \leq k} \frac{u_{k+1} \cdot v_i}{||v_i||^2} v_i$. While not necessary for the proof, we note that $u_{k+1} \notin span(\{u_1, \ldots, u_k\}) = span(\{v_1, \ldots, v_k\})$ so that $v_{k+1} \neq 0$.

We see that v_{k+1} can be expressed as a linear combination of u_{k+1} and $v_i, 1 \leq i \leq k$. Also, u_{k+1} can be expressed a linear combination of $v_i, 1 \leq i \leq k + 1$. Hence $span(\{v_1, \ldots, v_k, v_{k+1}\}) = span(\{v_1, \ldots, v_k, u_{k+1}\})$. Finally, by induction hypothesis $P(k)$, we replace $v_i, 1 \leq i \leq k$ by $u_i, 1 \leq i \leq k$ such that $span(\{v_1, \ldots, v_k, u_{k+1}\}) = span(\{u_1, \ldots, u_{k+1}\})$. Hence the induction holds.

At termination, $span(\{v_1, \ldots, v_n\}) = span(\{u_1, \ldots, u_n\})$ which means $v_i, 1 \leq i \leq n$ spans vector space V. Since there are $n = \dim(V)$ vectors, it is a basis.

If $n = 1$, the new basis is trivially orthogonal. If $n > 1$, we take any two distinct indices i, j in $\{1, \ldots, n\}$. WLOG, $i < j$, by the Gram-Schmidt process, v_j is the orthogonal component of u_j w.r.t. $span(\{v_1, \ldots, v_{j-1}\})$, hence in particular, $v_j \perp v_i$.

Use of CAS (Sagemath)   Suppose we wish to compute the solution to $A^T A x = b$, where $A = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 1 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$. This is based on example 5.3.11.3 of the text.

- `A = matrix([[1,1,2],[-1,2,1],[1,0,1],[-1,1,0]])`

- `b = vector([1,1,1,1])`

- To obtain RREF of the augmented matrix $[A^T A A^T b]$, `block_matrix([[A.transpose()*A, (A.transpose()*b).column()]]).rref()`

- After solving manually for a particular solution x, we define `u = vector(QQ, [2/5, 4/5, 0])`

- If we want to find Au, simply do `A*u`

Useful commands w.r.t. manipulating vectors and matrices. Suppose we define a vector `v = vector(...)`. Sagemath is intelligent enough to convert it to a column or row vector depending on how it is used, for e.g. premultiplying it to a matrix `v*A` causes it to be interpreted as a row vector.

To convert a vector to a row matrix, do `v.row()`.
To convert a vector to a column matrix, do `v.column()`.
The following are how Sagemath prints the vector, row matrix, col matrix respectively:

- v: $(v_1, v_2, \dots)$

- v.row(): $[v_1, v_2, \dots]$

- v.column(): $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix}$

See `wiki.sagemath.org/quickref?action=AttachFile&do=get&target=quickref-linalg.pdf` for complete reference.

Alternative proof to Theorem 5.4.7   Claim: Let $S = \{u_1, u_2, \dots, u_n\}, T = \{v_1, v_2, \dots, v_n\}$ be 2 orthonormal bases for vsp V. Then the transition matrix $P_{S \to T}$ is orthogonal.

In the text, both $P_{S \to T}$ and $P_{T \to S}$ are explicitly evaluated and found to be transposes. Instead of this, we make the observation that $P_{S \to T} = B^{-1}A$, where $A = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}, B = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$.

Since S, T are orthonormal, the columns of A and B are orthonormal, so A, B are orthogonal. Since orthogonal matrices are closed under product and inverses, $P_{S \to T} = B^{-1}A$ is also orthogonal.

We remark that this proof says that the transition matrix being orthogonal is just a simple consequence of closure properties of the product of orthogonal matrices (in particular the orthogonal group of $n \times n$ real matrices $O(n, \mathbb{R})$).

Uniqueness of standard matrix   In the text, a linear map $T \in L(\mathbb{R}^n, \mathbb{R}^m)$ is defined as a function of the form $v \mapsto Av$, where $A \in \mathbb{R}^{m,n}$ is known as the standard matrix of T.

A trivial but natural question would be the well-definition of A. And we can indeed see that this is the case. Suppose A_1, A_2 are 2 matrices that are both standard matrices of the same linear map T. Then $T = v \mapsto A_1 v \wedge T = v \mapsto A_2 v$, which says $T - T = v \mapsto (A_1 - A_2)v$ is the zero function. In particular, for canonical basis vectors $e_j, 1 \le j \le n$, $(v \mapsto (A_1 - A_2)v)(e_j) = (A_1 - A_2)e_j = \mathbf{0}$. Letting $\epsilon_i, 1 \le i \le m$ be the ith projection function, we have $(A_1 - A_2)_{i,j} = \epsilon_i((A_1 - A_2)e_j) = \epsilon_i(\mathbf{0}) = 0$, hence $A_1 - A_2$ is the zero matrix, which says $A_1 = A_2$. Hence the standard matrix is unique.

Geometrical transformations   A linear map T's action on a vector space V is completely determined by T's action on any basis of V due to linearity property.

Hence, to determine the transformation matrix of a certain geometrical LT in $\mathbb{R}^2$ or $\mathbb{R}^3$, it suffices to consider how the LT acts on the standard basis. i.e. the transformation matrix is $[T(e_1)T(e_2)]$ (for 2-space).

Homogeneous coordinates   This topic belongs to the last subsection of the text, which is not covered in the module. However, the part about homogeneous coordinates is worth the read.

Amongst the elementary geometrical transformations (scaling, translations, rotations, shearing), all are linear maps except translations. Translations are handled by increasing the dimension by 1 and making use of a shearing in the extra coordinate.

The "true" coordinates of a point ignores the extra coordinate (for e.g. (x, y, w) in 2-space really refers to the point (x, y) and (x, y, z, w) in 3-space refers to the point (x, y, z)), as it is only there for the purposes of conducting transformations.