Summary of week 11 (Lectures 31, 32 and 33)

The contents of this week's lectures coincided almost exactly with Chapter 7 of [VST]. This involved repeating some of the material from Week 5.

Recall that if V is an n-dimensional vector space over F and $\boldsymbol{b}$ is a basis of V consisting of the vectors v_1, v_2, $\ldots$, v_n then there is an isomorphism $\phi\colon F^n \to V$ given by

$$\phi\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix} = \sum_{i=1} \lambda_i v_i.$$

Let $\mathrm{cv}_{\boldsymbol{b}}\colon V \to F^n$ be the isomorphism that is the inverse of ϕ. In other words, for all $v \in V$,

$$\mathrm{cv}_{\boldsymbol{b}}(v) = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix} \in F^n$$

where the λ_i are the unique scalars such that $v = \sum_{i=1}^{n} \lambda_i v_i$. The n-tuple $\mathrm{cv}_{\boldsymbol{b}}(v)$ is called the coordinate vector of v relative to the basis $\boldsymbol{b}$.

Suppose we also have an m-dimensional space W with basis $\boldsymbol{c}$ consisting of w_1, w_2, $\ldots$, w_m, and a linear map $T\colon V \to W$. Just as for V we have mutually inverse isomorphisms $\psi\colon F^m \to W$ and $\mathrm{cv}_{\boldsymbol{c}}\, W\colon F^m$ such that for all $w \in W$,

$$\mathrm{cv}_{\boldsymbol{c}}(w) = \begin{pmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_m\end{pmatrix} \quad \text{if and only if} \quad w = \sum_{i=1}^{m} \mu_i w_i.$$

We define the *matrix of T relative to the bases $\boldsymbol{c}$ and $\boldsymbol{b}$* to be the $m \times n$ matrix $\mathrm{M}_{\boldsymbol{cb}}(T)$ whose j-th column is $\mathrm{cv}_{\boldsymbol{c}}(T(v_j))$. That is, if a_{ij} is the (i,j)-entry of $\mathrm{M}_{\boldsymbol{cb}}(T)$, then $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$.

Now for arbitrary $v \in V$, if we let

$$\mathrm{cv}_{\boldsymbol{b}}(v) = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix},$$

then we find that

$$\begin{aligned} T(v) = T\Big(\sum_{j=1}^{n} \lambda_j v_j\Big) &= \sum_{j=1}^{n} \lambda_j T(v_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} w_i \\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_j a_{ij} w_i = \sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_j a_{ij} w_i = \sum_{i=1}^{m}\Big(\sum_{j=1}^{n} a_{ij}\lambda_j\Big) w_i, \end{aligned}$$

and consequently the coordinate vector of $T(v)$ relative to $\boldsymbol{c}$ is given by

$$\mathrm{cv}_{\boldsymbol{c}}(T(v)) = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\lambda_j \\ \sum_{j=1}^{n} a_{2j}\lambda_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}\lambda_j \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \mathrm{M}_{\boldsymbol{cb}}(T)\,\mathrm{cv}_{\boldsymbol{b}}(v).$$

This result is Theorem 7.1 of [VST].

We can also prove Theorem 7.1 by the following argument. We proved last week that composites of linear maps are linear. Hence, forming the composite of the three linear maps $\phi\colon F^n \to V$, followed by $T\colon V \to W$, followed by $\mathrm{cv}_{\boldsymbol{c}}\colon W \to F^m$, we obtain a linear map $F^n \to F^m$. As the diagram below indicates, this map takes $\mathrm{cv}_{\boldsymbol{b}}(v)$ to $\mathrm{cv}_{\boldsymbol{c}}(T(v))$ (for all $v \in V$). But Proposition 7.2 of

$$\begin{tikzcd}
V & W \\
v = \phi(\mathrm{cv}_{\boldsymbol{b}}(v)) \arrow[r, mapsto, "T"] & T(v) \arrow[d, mapsto, "\mathrm{cv}_{\boldsymbol{c}}"] \\
\mathrm{cv}_{\boldsymbol{b}}(v) \arrow[u, mapsto, "\phi"] & \mathrm{cv}_{\boldsymbol{c}}(T(v)) \\
F^n & F^m
\end{tikzcd}$$

[VST], which we proved by a direct argument in Week 5, tell us that every linear map $F^n \to F^m$ is given by multiplication by an $m \times n$ matrix; so there exists $A \in \mathrm{Mat}(m \times n, F)$ such that $A\,\mathrm{cv}_{\boldsymbol{b}}(v) = \mathrm{cv}_{\boldsymbol{c}}(T(v))$ for all $v \in V$. The matrix A is $\mathrm{M}_{\boldsymbol{cb}}(T)$, the matrix of T relative to the given bases.

Choosing a basis for a vector space can be thought of as choosing a coordinate system, and we can think of $\mathrm{M}_{\boldsymbol{cb}}(T)$ as the coordinate form of T, relative to the chosen coordinate systems, in the same way that $\mathrm{cv}_{\boldsymbol{b}}(v)$ and $\mathrm{cv}_{\boldsymbol{c}}(T(v))$ are the coordinates of v and $T(v)$.

Here is the rule you should remember.

To calculate the matrix of a linear map

Given a linear map $T\colon V \to W$ and bases for V and W, apply T to the first vector in the basis of V and express the result as a linear combination of the basis elements of W. The coefficients obtained form the first column of the required matrix. Repeat for the second basis vector of V to get the second column, and so on for all the elements in the basis of V.

For example, let V be the vector space over $\mathbb{R}$ consisting of all polynomials in the variable x of degree at most three, and let W consist of all polynomials of degree at most 2. Differentiation defines a map D from V to W, and it is immediate from basic rules of calculus that D is linear. (Specifically, the rules in question say that $\frac{d}{dx}(p(x)+q(x)) = \frac{d}{dx}p(x) + \frac{d}{dx}q(x)$, and $\frac{d}{dx}(\lambda p(x)) = \lambda \frac{d}{dx}p(x)$.) Let $\boldsymbol{b}$ be the basis of V consisting of the polynomials x^3, x^2, x and 1, and let $\boldsymbol{c}$ be the basis of W consisting of x^2, x and 1. Now

$$\begin{aligned} D(x^3) &= 3x^2 = 3x^2 + 0x + 0, \qquad & D(x) &= 1 = 0x^2 + 0x + 1, \\ D(x^2) &= 2x = 0x^2 + 2x + 0, & D(1) &= 0 = 0x^2 + 0x + 0, \end{aligned}$$

and so by the rule stated above it follows that

$$\mathrm{M}_{\boldsymbol{cb}}(D) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

It is trivial to confirm that Theorem 7.1 holds in this case. If $p(x)$ is an arbitrary element of V then $p(x) = ax^3 + bx^2 + cx + d$ for some scalars a, b, c, $d \in \mathbb{R}$, and we have

$$\mathrm{cv}_{\boldsymbol{b}}(p(x)) = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

So according to Theorem 7.1 we should have that

$$\mathrm{cv}_{\boldsymbol{c}}(D(p(x))) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3a \\ 2b \\ c \end{pmatrix}.$$

This tells us how to express $D(p(x))$ as a linear combination of the elements of the basis $\boldsymbol{c}$; specifically, it says that

$$D(p(x)) = 3ax^2 + 2bx + c,$$

which is, of course, precisely what one gets by differentiating $p(x)$ in the usual way.

Here is another example. Let $\boldsymbol{b}$ be the basis of $\mathbb{R}^3$ consisting of

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

and let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+2y+z \\ 3x+z \\ x-y-z \end{pmatrix}.$$

It is easy to check that T is a linear transformation. We proceed to calculate $\mathrm{M}_{\boldsymbol{bb}}(T)$, the matrix of T relative to the basis $\boldsymbol{b}$ of the domain of T and the basis $\boldsymbol{b}$ of the codomain of T. Of course in this example the domain of T and the codomain of T are equal; so we have the possibility of using the same basis for them both, and in this case we have chosen to do so. On other occasions we may use different bases even when the domain and codomain are the same.

Calculation gives

$$\begin{aligned} T(v_1) = T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\ T(v_2) = T\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\ T(v_3) = T\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} &= \begin{pmatrix} 4 \\ 4 \\ -1 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 5\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 1\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \end{aligned}$$

and we deduce that

$$\mathrm{M}_{\boldsymbol{bb}}(T) = \begin{pmatrix} -2 & 0 & 0 \\ 2 & 3 & 5 \\ 1 & 0 & -1 \end{pmatrix}.$$

In confirmation of Theorem 7.1, observe that since

$$\begin{pmatrix} -2 & 0 & 0 \\ 2 & 3 & 5 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 23 \\ -2 \end{pmatrix}$$

it should be the case that T takes the vector

$$1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \\ 3 \end{pmatrix}$$

to the vector

$$-2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 23\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 19 \\ 21 \\ -2 \end{pmatrix}.$$

This is left for the reader to check.

Theorem 4.16 of [VST] tells us that if $v_1, v_2, \ldots, v_n$ is a basis of V and $w_1, w_2, \ldots, w_m$ a basis of W then there is a linear map $V \to W$ taking the v_i to any arbitrarily chosen linear combinations of the w_j; so any $m \times n$ matrix is the matrix relative to these two bases of some linear map $V \to W$. For example, if $n = 3$ and $m = 4$ then there is a linear map ϕ such that

$$\begin{aligned}\phi(v_1) &= 2w_1 - w_2 + 4w_3 - w_4\\ \phi(v_2) &= w_1 + w_2 - w_3 + 2w_4\\ \phi(v_3) &= -w_1 + w_2 + 0w_3 - 2w_4,\end{aligned}$$

and its matrix is

$$\begin{pmatrix} 2 & 1 & -1\\ -1 & 1 & 1\\ 4 & -1 & 0\\ -1 & 2 & -2\end{pmatrix}.$$

Let V and U be arbitrary finite-dimensional vector spaces and $\phi: V \to U$ an arbitrary linear map. As in our discussion of the Main Theorem on Linear Transformations, let $K = \ker\phi$ (a subspace of V) and $I = \operatorname{im}\phi$ (a subspace of W), and choose a subspace W of V that is complementary to K. The proof of the Main Theorem consisted of showing that W is isomorphic to I. More specifically, there is an isomorphism $W \to I$ such that $w \mapsto \phi(w)$ for all $w \in W$.

Now let $x_1, x_2, \ldots, x_r$ be a basis of W, and let $x_{r+1}, x_{r+2}, \ldots, x_n$ be a basis of K. Since $V = W \oplus K$, combining these bases gives a basis $x_1, x_2, \ldots, x_n$ of V (by Theorem 6.9 of [VST]). Call this basis $\boldsymbol{b}$.

For each $i \in \{1, 2, \ldots, r\}$ let $y_i = \theta(x_i) \in I$. Since the isomorphism $W \to I$ described above takes $x_i \mapsto y_i$, it follows from Theorem 4.17 that $y_1, y_2, \ldots, y_r$ is a basis of I. By Proposition 4.10 we can extend this basis to a basis $y_1, y_2, \ldots, y_m$ of U. Call this basis $\boldsymbol{c}$. Now we find that

$$\begin{aligned}\phi(x_1) &= 1y_1 + 0y_2 + \cdots + 0y_r + 0y_{r+1} + \cdots + 0y_m\\ \phi(x_2) &= 0y_1 + 1y_2 + \cdots + 0y_r + 0y_{r+1} + \cdots + 0y_m\\ &\ \vdots\\ \phi(x_r) &= 0y_1 + 0y_2 + \cdots + 1y_r + 0y_{r+1} + \cdots + 0y_m\\ \phi(x_{r+1}) &= 0y_1 + 0y_2 + \cdots + 0y_r + 0y_{r+1} + \cdots + 0y_m\\ &\ \vdots\\ \phi(x_r) &= 0y_1 + 0y_2 + \cdots + 0y_r + 0y_{r+1} + \cdots + 0y_m\end{aligned}$$

where the first r of these equations come from the definition of the elements y_i while the remaining $n - r$ equations (which say that $\phi(x_j) = 0$ for $r + 1 \leq j \leq n$) come from the fact the $x_{r+1}, x_{r+2}, \ldots, x_m \in K = \ker\phi$. The coefficients on the right hand side of the first of these equations gives the entries in the first column of $\mathrm{M}_{\boldsymbol{cb}}(\phi)$, the second column is similarly obtained from the second equation, and so on. So

$$\mathrm{M}_{\boldsymbol{cb}}(\phi) = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots\dots\dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots\dots\dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots\dots\dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots\dots\dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots\dots\dots & 0 \end{pmatrix}.$$

We have shown that if we choose bases to suit a given linear map then we can arrange for the matrix of the map to have the simple form that appears above. Note that the number of 1's in the matrix is the dimension of the image of ϕ. We call this number the *rank* of ϕ.

The result proved above is Theorem 7.12 of [VST].

Since it is possible to so dramatically simplify the matrix of a linear map by changing bases, it is important for us to understand precisely how a change of basis alters the matrix. Here is the key definition.

Definition. Let V be a finite-dimensional vector space and let $\boldsymbol{b}$ and $\boldsymbol{c}$ be bases of V. Let $\mathrm{i}\colon V \to V$ be the identity map. The *transition matrix from* $\boldsymbol{b}$*-coordinates to* $\boldsymbol{c}$*-coordinates* is the matrix $\mathrm{M}_{\boldsymbol{cb}}$ defined by

$$\mathrm{M}_{\boldsymbol{cb}} = \mathrm{M}_{\boldsymbol{cb}}(\mathrm{i}).$$

Observe that, in the situation of the above definition, if $v \in V$ is arbitrary then

$$\mathrm{M}_{\boldsymbol{cb}}\,\mathrm{cv}_{\boldsymbol{b}}(v) = \mathrm{M}_{\boldsymbol{cb}}(\mathrm{i})\,\mathrm{cv}_{\boldsymbol{b}}(v) = \mathrm{cv}_{\boldsymbol{c}}(\mathrm{i}(v)) = \mathrm{cv}_{\boldsymbol{c}}(v).$$

That is, multiplication by the transition matrix takes the coordinate vector of v relative to $\boldsymbol{b}$ to the coordinate vector of v relative to $\boldsymbol{c}$.

As an example, consider the following two bases of $\mathbb{R}^2$:

$$\boldsymbol{b} = \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right),$$
$$\boldsymbol{c} = \left(\begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right).$$

To find $\mathrm{M}_{\boldsymbol{cb}}$ we express the elements of $\boldsymbol{b}$ as linear combinations of the elements of $\boldsymbol{c}$. It is readily found that

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$
$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 5 \end{pmatrix} - 3\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

and therefore $\mathrm{M}_{\boldsymbol{cb}} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}$. Now, for example,

$$8\begin{pmatrix} 1 \\ 2 \end{pmatrix} - 10\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \end{pmatrix}$$

and so $\mathrm{cv}_{\boldsymbol{b}}\begin{pmatrix} -2 \\ 6 \end{pmatrix} = \begin{pmatrix} 8 \\ -10 \end{pmatrix}$. It now follows that

$$\mathrm{cv}_{\boldsymbol{c}}\begin{pmatrix} -2 \\ 6 \end{pmatrix} = \mathrm{M}_{\boldsymbol{cb}}\begin{pmatrix} 8 \\ -10 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 8 \\ -10 \end{pmatrix} = \begin{pmatrix} -12 \\ 22 \end{pmatrix}.$$

Direct computation confirms it to be true that $\begin{pmatrix} -2 \\ 6 \end{pmatrix} = -12\begin{pmatrix} 2 \\ 5 \end{pmatrix} + 22\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

The next ingredient that we need is Theorem 7.5 of [VST]; it say that composition of linear maps agrees with multiplication of their matrices. To be precise, suppose that $\phi\colon U \to V$ and $\psi\colon V \to W$ are linear maps, and let

$$X = \mathrm{M}_{\boldsymbol{cb}}(\phi)$$
$$Y = \mathrm{M}_{\boldsymbol{dc}}(\psi)$$

where $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$ are bases of U, V, W respectively. Then

$$YX = \mathrm{M}_{\boldsymbol{db}}(\psi\phi).$$

See pages 154 and 155 of [VST] for the proof.

It is a triviality that for any basis $\boldsymbol{b}$ of any space V, if i is the identity map $V \to V$ then $\mathrm{M}_{\boldsymbol{bb}}(\mathrm{i})$ is the identity matrix I. We readily deduce Corollary 7.6 of [VST]: if ϕ is a vector space isomorphism then the matrix of the inverse of ϕ is the inverse of the matrix of ϕ. That is, $\mathrm{M}_{\boldsymbol{bc}}(\phi^{-1}) = \mathrm{M}_{\boldsymbol{cb}}(\phi)^{-1}$. In the special case that ϕ is the identity map on a space V this says that the transition matrix from $\boldsymbol{c}$-coordinates to $\boldsymbol{b}$-coordinates is the inverse of the transition matrix from $\boldsymbol{c}$-coordinates to $\boldsymbol{b}$-coordinates.

Another important corollary of Theorem 7.5 answers the question of how changing bases affects the matrix of a linear map: if $\phi\colon V \to W$ is linear and $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ are bases of V and $\boldsymbol{c}_1$, $\boldsymbol{c}_2$ bases of W, then

$$\mathrm{M}_{\boldsymbol{c_2 b_2}}(\phi) = \mathrm{M}_{\boldsymbol{c_2 c_1}}\,\mathrm{M}_{\boldsymbol{c_1 b_1}}(\phi)\,\mathrm{M}_{\boldsymbol{b_1 b_2}}\,.$$

Lecture 33 included a discussion of elementary column operations. Since transposing matrices changes columns into rows it is not surprising that for every fact about row operations there is a corresponding fact about column operations. Since transposing reverses matrix multiplication, it turns out that postmultiplying a matrix A by an elementary matrix performs a column operation on A, whereas premultiplying A by an elementary matrix performs a row operation on A. See pages 33 and 34 of [VST] for more details.

Two matrices are said to be *row-equivalent* if it is possible to transform one into the other by performing a sequence of elementary row operations. Similarly, two matrices are said to be *column-equivalent* if it is possible to transform one into the other by performing a sequence of elementary column operations. Thus A and B are column-equivalent if and only if $B = AE_1E_2 \cdots E_k$ for some elementary matrices E_1, E_2, …, E_k. Since a matrix is invertible if and only if it can be expressed as a product of elementary matrices, we see that A and B are column-equivalent if and only if $B = AP$ for some invertible P. Analogously, A and B are row-equivalent if and only if $B = PA$ for some invertible P.

Recall that the column space of a matrix $A \in \mathrm{Mat}(m \times n, F)$ coincides with the set $\{\, Ay \mid y \in F^n \,\}$. If $B = AP$ for some $P \in \mathrm{Mat}(n \times n, F)$ then

$$\begin{aligned} \mathrm{CS}(B) &= \{\, Bx \mid x \in F^n \,\} \\ &= \{\, A(Px) \mid x \in F^n \,\} \\ &\subseteq \{\, Ay \mid y \in F^n \,\} = \mathrm{CS}(A). \end{aligned}$$

If P is invertible then $A = BP^{-1}$, and so we deduce also that $\mathrm{CS}(A) \subseteq \mathrm{CS}(B)$. Hence it follows that if A and B are column equivalent then they have the same column space. Of course we also have the corresponding result for rows: if A and B are row-equivalent then they have the same row space. See 3.21 and 3.22 (pp. 74–75) of [VST] for more details.

Performing an elementary column operation on a matrix produces a column equivalent matrix, and so does not change the column space. Performing an elementary row operation is, of course, likely to change the column space. However, it is not hard to see that at least the dimension of the column space is not changed by row operations. Indeed, if $B = PA$, where P is invertible, then premultiplying any element of the set $Ax \mid x \in F^n$ by the matrix P will give an element of the set $PAx \mid x \in F^n = Bx \mid x \in F^n$. In other words, premultiplication by P defines a map from $\mathrm{CS}(A)$ to $\mathrm{CS}(B)$. Similarly, premultiplication by P^{-1} maps $\mathrm{CS}(B)$ to $\mathrm{CS}(A)$. These maps are both linear and are inverse to each other; hence they are isomorphisms, and hence they preserve dimension, as claimed.

Finally, Theorem 7.22 of [VST] was proved. This result says that an arbitrary $m \times n$ matrix A is equivalent to one which has 1's in the $(1,1)$, $(2,2)$, …, (r,r) positions, for some r, and zeros elsewhere. Here the word "equivalent" is used in the following sense: A is equivalent to B if it is possible to transform A into B by using an arbitrary combination of row and column operations. This is the same as saying that $B = PAQ$ for some invertible P and Q. Since neither the dimension of the row space nor the dimension of the column space changes under elementary row or column operations it follows that the number r is both the dimension of the row space and the dimension of the column space. This number is called the *rank* of the matrix A.