Inverting non-square matrices v 2 PhL 3.3.16

I think my original document is very incomplete and I need to go read about this topic in some good exposition.

My matrix binder deals only with square matrices!!!

What books do I already have? Well, I am pretty sure that no physical book I have talks about this subject! It is really a Linear Algebra topic and I have no physical books on LA.

What about pdf books in stock? I have a LA section both in math and in mathdownloads. In the first I have

Toomas: (188 pages) (no TOC, no index, yes OCR). I see nothing paging through on non-square.

Matthews: only PDF chunks, does not seem useful

Allen: Separate PDF's for chapters. TOC does not seem to show what I want.

In the downloads area I have

Allen all in one: nada

Beezer: mentions left and right inverses only in a few places, no go

Hefferon: talks about left inverses of maps, but I don't really see what I want in notation I want.

Horn: nada

Schaum: nada on left inverse

Only Shilov has anything!

Shilov: he *does* have some stuff. For example on page 97

THEOREM. *The operator* $\mathbf{A}:\mathbf{X}\to\mathbf{Y}$ *has a left inverse if and only if* $\mathbf{A}$ *is a monomorphism. The operator* $\mathbf{B}:\mathbf{Y}\to\mathbf{X}$ *has a right inverse if and only if* $\mathbf{B}$ *is an epimorphism.*

Now you have to go look up those words:

2.7. Morphisms of Linear Spaces

2.71. Let ω be a rule which assigns to every given vector x' of a linear space $\mathbf{K}'$ a vector x'' in a linear space $\mathbf{K}''$. Then ω is called a *morphism* (or *linear operator*)† if the following two conditions hold:

a) $\omega(x' + y') = \omega(x') + \omega(y')$ for every $x', y' \in \mathbf{K}'$;
b) $\omega(\alpha x') = \alpha\omega(x')$ for every $x' \in \mathbf{K}'$ and every $\alpha \in K$.

A morphism ω mapping the space $\mathbf{K}'$ onto the whole space $\mathbf{K}''$ is called an *epimorphism*. A morphism ω mapping $\mathbf{K}'$ onto part (or all) of $\mathbf{K}''$ in a one-to-one fashion (so that $x' \neq y'$ implies $\omega(x') \neq \omega(y')$) is called a *monomorphism*. A morphism ω mapping $\mathbf{K}'$ onto all of $\mathbf{K}''$ in a one-to-one fashion (i.e., a morphism which is both an epimorphism and a monomorphism) is called an *isomorphism*, and the spaces $\mathbf{K}'$ and $\mathbf{K}''$ themselves are said to be *isomorphic* (more exactly, *K-isomorphic*). The usual notation for a morphism is

$$\omega : \mathbf{K}' \to \mathbf{K}''.$$

† More exactly, a *morphism of* $\mathbf{K}'$ *into* $\mathbf{K}''$ (or a *linear operator mapping* $\mathbf{K}'$ *into* $\mathbf{K}''$).

I would say that morphism = linear operator, find. I think epimorphism means onto and also surjection, confirmed by web. Now monomorphism means 1-to-1 between K' and part or all of K" as range. If all of K" is involved you have an isomorphism. Fine! So he is saying above

In AX = Y, A has a left inverse iff this is 1-to-1 between domain and range. Very reasonable. It is then just the obvious "inverse mapping".

In BY = X suppose B is onto. Then for every X in the entire range space there must some Y in the domain. Not sure why this is called a right inverse. I would need a left inverse to write $Y = B^{-1}X$.

THEOREM. *An $m \times n$ matrix A has a left inverse if and only if its rank equals n and a right inverse if and only if its rank equals m.*

As usual in Shilov, there is no proof, you are referred to a scattered tree of past results. I started into this below but then gave up parsing. It is the above result that I really wanted to see.

****

So there you are, a well stated theorem, something I wanted to see. I think the proof is on page 94. Let's see if I can follow it, clip at a time. First we need this definition

The coefficients $a_i^{(j)}$ $(i = 1, \ldots, m; j = 1, \ldots, n)$ define an $m \times n$ matrix†

$$A = A_{(e,f)} = \begin{Vmatrix} a_1^{(1)} & a_1^{(2)} & \cdots & a_1^{(n)} \\ a_2^{(1)} & a_2^{(2)} & \cdots & a_2^{(n)} \\ \cdot & \cdot & \cdots & \cdot \\ a_m^{(1)} & a_m^{(2)} & \cdots & a_m^{(n)} \end{Vmatrix},$$

so for me this says

$A_i{}^j = a_i{}^{(j)}$ Now we start:

4.62. Next let $\mathbf{N}(\mathbf{A})$ be the *null space* of the operator $\mathbf{A}$, i.e., the set of all vectors $x \in \mathbf{X}$ such that $\mathbf{A}x = 0$, and as before let $A = \|a_i^{(j)}\|$ be the matrix of $\mathbf{A}$. We now consider the problem of finding the dimension of the subspace $\mathbf{N}(\mathbf{A})$ from a knowledge of the matrix A. Let

$$x = \sum_{i=1}^{n} \xi_i e_i \in \mathbf{N}(\mathbf{A}).$$

Then the system (5), p. 80 takes the form

$$\begin{aligned} a_1^{(1)}\xi_1 + a_1^{(2)}\xi_2 + \cdots + a_1^{(n)}\xi_n &= 0, \\ a_2^{(1)}\xi_1 + a_2^{(2)}\xi_2 + \cdots + a_2^{(n)}\xi_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_m^{(1)}\xi_1 + a_m^{(2)}\xi_2 + \cdots + a_m^{(n)}\xi_n &= 0. \end{aligned} \qquad (16)$$

OK, he is going to focus on the nullspace of A, and for some strange reason he writes $A_i{}^j$ as $a_i{}^{(j)}$, fine. The nullspace means A**x** = 0 which then means

$A\,[\,\Sigma_i \xi_i \mathbf{e}_i] = 0 \quad \Rightarrow \quad \Sigma_i \xi_i\, A\mathbf{e}_i = 0 \quad \Rightarrow \Sigma_i \xi_i\, [A\mathbf{e}_i]_k = 0$

$\Rightarrow \quad \Sigma_i \xi_i\, \Sigma_s\, A_{ks}\, (\mathbf{e}_i)_s = 0 \quad \Rightarrow \Sigma_s\, \Sigma_i \xi_i\, A_{ks}\, \delta_{i,s} = 0$

$\Rightarrow \quad \Sigma_s\, \xi_s\, A_{ks} = 0$

Write this out to get

k = 1 $\quad \xi_1 A_{11} + \xi_2 A_{12} + \ldots = 0$

k = 2 $\quad \xi_1 A_{21} + \xi_2 A_{22} + \ldots = 0$

which I rewrite as

k = 1 $\quad a_1{}^{(1)}\xi_1 + a_1{}^{(2)}\xi_2 + \ldots = 0$

k = 2 $\quad a_2{}^{(1)}\xi_1 + a_2{}^{(2)}\xi_2 + \ldots = 0$

So I agree with (16) above. This involves the coefficients ξ_i and assumes $(\mathbf{e}_i)_s = \delta_{i,s}$ so this is a certain standard basis I use but here not in covariant notation. The equations involve matrix elements of A and the coefficients. All just fine, so go for the next clip:

Therefore from a knowledge of the matrix of the operator **A** relative to the basis $e_1, e_2, \ldots, e_n$ we can determine the result of applying **A** to any vector

$$x = \sum_{j=1}^{n} \xi_j e_j$$

of the space **X**. In fact, the equations (5) express the components of the vector $y = \mathbf{A}x$ as linear combinations of the components of x. Note that *the coefficient matrix of the system of the equations* (5) *is just the matrix* $A_{(e,f)}$.

This does not seem to add anything new, just a restatement of what we just did. Continue,

Next let $\|a_i^{(j)}\|$ be an arbitrary $m \times n$ matrix, where the superscript is the column number and the subscript is the row number. Given any vector

$$x = \sum_{j=1}^{n} \xi_j e_j,$$

we construct the vector

$$y = \sum_{i=1}^{m} \eta_i f_i$$

with components $\eta_1, \eta_2, \ldots, \eta_m$ determined by (5). It is easy to see that

But what are the f_i here? Have to go back. Our context is A: X → Y where X and Y are spaces of different dimension if A is non-square. Back up:

4.23. Matrix representation of linear operators. Let **A** be a linear operator mapping a space **X** of dimension n into a space **Y** of dimension m. Let $e_1, \ldots, e_n$ be a fixed basis in **X** and $f_1, \ldots, f_m$ a fixed basis in **Y**. The vector e_1 is mapped by **A** into some vector $\mathbf{A}e_1$ of the space **Y**, which, like every vector of **Y**, has an expansion

$$\mathbf{A}e_1 = a_1^{(1)} f_1 + a_2^{(1)} f_2 + \cdots + a_m^{(1)} f_m$$

with respect to the basis vectors $f_1, \ldots, f_m$. The operator **A** has a similar effect on the other basis vectors:

$$\mathbf{A}e_2 = a_1^{(2)} f_1 + a_2^{(2)} f_2 + \cdots + a_m^{(2)} f_m,$$

OK, so in our clip above we have A**x** = **y** where **x** is in x-space of dimension n and basis $\mathbf{e_i}$ and where **y** is in y-space of dimension m and basis $\mathbf{f_i}$. So the above expansions of **x** and **y** are fine by me and we can continue to another clip:

with components $\eta_1, \eta_2, \ldots, \eta_m$ determined by (5). It is easy to see that the operator **A** effecting this mapping of the vector x into the vector y is a linear operator. We now construct the matrix of the operator **A** relative to the basis $e_1, e_2, \ldots, e_n$. Since the vector e_1 has components $\xi_1 = 1$, $\xi_2 = 0$, $\ldots, \xi_n = 0$, it follows from (5) that the components of the vector $\mathbf{A}e_1$ will be the numbers $a_1^{(1)}, a_2^{(1)}, \ldots, a_m^{(1)}$, so that

$$\mathbf{A}e_1 = a_1^{(1)} f_1 + a_2^{(1)} f_2 + \cdots + a_m^{(1)} f_m.$$

Similarly,

$$\mathbf{A}e_j = a_1^{(j)} f_1 + a_2^{(j)} f_2 + \cdots + a_m^{(j)} f_m \qquad (j = 1, 2, \ldots, n).$$

OK, here he confirms my understanding of his basis vectors $\mathbf{e_i}$ But then what he next says seems illogical to me. I agree that $\mathbf{e_1}$ = (1,0,0...) as he says. Now here is equation (5) for **y** = A**x**

$$\begin{aligned}
\eta_1 &= a_1^{(1)}\xi_1 + a_1^{(2)}\xi_2 + \cdots + a_1^{(n)}\xi_n, \\
\eta_2 &= a_2^{(1)}\xi_1 + a_2^{(2)}\xi_2 + \cdots + a_2^{(n)}\xi_n, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\eta_m &= a_m^{(1)}\xi_1 + a_m^{(2)}\xi_2 + \cdots + a_m^{(n)}\xi_n.
\end{aligned} \qquad (5)$$

where η_i are the components of **y** in its appropriate space. So for $\mathbf{e_1}$ we get

$\eta_1 = a_1{}^{(1)}$
$\eta_2 = a_2{}^{(1)}$ etc

so then I agree

$\mathbf{A}e_1 = \eta_1\mathbf{f_1} + \eta_2\mathbf{f_2} + ... = a_1{}^{(1)}\mathbf{f_1} + a_2{}^{(1)}\mathbf{f_2} + ...$

so I agree with the first of his equations, and the generalization seems obvious as his last line above. Now on to the next clip: