Notes on Bonus Questions for MAT2611

Partha Pratim Ghosh

In this document I shall provide you with all the possible hints for the Bonus Questions. These are not complete solutions. These are hints, remarks about how to solve the problem, ... The solutions have to be completed by yourselves only! Remember: you learn to *swim* only by *swimming*.

Please feel free to send your further queries directly to me at my email address ghoshpp@unisa.ac.za. Please **do not** send your emails through the *myUnisa* website or the Microsoft Office 365 service, since there are complaints of students finding me unreachable from any of the above two services; use an independent email client instead.

1. Video Lecture 1

Problem 1. For each of the following show that they are indeed examples of a vector space:

(a) Given any $n \geq 1$ the set $\mathbb{R}^n$ with the addition and scalar multiplication defined by:

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

and

$$\lambda(x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n).$$

(b) Given any set X the set $\mathbb{R}^X$ of all real valued functions on X with vector addition and scalar multiplication defined by:

$$(f + g)(x) = f(x) + g(x), \quad \text{for all } x \in X,$$

and

$$(\lambda f)(x) = \lambda f(x),$$

where $X \underset{g}{\overset{f}{\rightrightarrows}} \mathbb{R}$ are two functions and $\lambda \in \mathbb{R}$.

(c) The set $C([0,1])$ of all real valued continuous functions on $[0,1]$ with usual addition and scalar multiplication of functions.

(d) The set $C_1([0,1])$ of all real valued functions on $[0,1]$ which are differentiable with usual addition and scalar multiplication of functions.

(e) The set $\mathbb{R}[x]$ of all polynomial functions with real coefficients with usual addition and scalar multiplication of polynomials.

(f) Given a natural number $n \geq 1$, the set $\mathbb{R}[x]_{\deg \leq n}$ of polynomial functions of degree at most n with usual addition and scalar multiplication of polynomials.

[$6 \times 10 = 60$ marks]

Sketches of Solution 1. (a) The closure properties for the vector addition and scalar multiplication for a vector space V are not necessary to check, if you have already established that the operations $V \times V \xrightarrow{+} V$ and $\mathbb{R} \times V \to V$ are defined. It is exactly in this statement that the statement of the closure axioms are encoded — for instance, if you say that $V \times V \xrightarrow{+} V$ is a function then you are actually saying that for each pair (u, v) of vectors from V, $+(u, v) = u + v \in V$, i.e. is a vector from V, which is actually the closure for vector addition.

Sometimes, the given formula for the addition or scalar multiplication may not be a function — for instance in Problem 2(c), the scalar multiplication is just not defined!

Thus, once you assert that $V \times V \xrightarrow{+} V$ and $V \xrightarrow{\mathbb{R} \times V}$ are well defined functions then you need not check the closure axioms.

(b) Just the closure axioms are not enough to grant a vector space.

For instance, consider the set $\mathbb{R}$ of real numbers with $\mathbb{R} \times \mathbb{R} \xrightarrow{+} \mathbb{R}$ defined now by $(x, y) \mapsto x - y$ and scalar multiplication being the usual multiplication of real numbers.

Surely, with these operations the closure properties are satisfied, but the operation $(x, y) \mapsto x - y$ is not commutative, not associative and does not have an identity! So it is not a vector space.

(c) One has to be careful with the subtle difference between a *function* and its *values*.

Recall that a function $X \xrightarrow{f} Y$ is a *rule* which assigns to each $x \in X$ a unique element $y \in Y$, and we write this unique $y \in X$ as $f(x)$. The alternative notations are: $y = f(x)$, $x \mapsto f(x)$, $x \overset{f}{\longmapsto} f(x)\ldots$. The element $f(x) \in Y$ is the *value* of the function f at the element $x \in X$ and f is not any of its *values*.

If $X \underset{g}{\overset{f}{\rightrightarrows}} Y$ are two functions between the same sets then, $f = g$, if and only if, for each $x \in X$, $f(x) = g(x)$.

So, in the context of (b), how do you show that $f + g$ is commutative? You should do something like — since f, g and $f + g$ have the same domain and codomain, for each $x \in X$, $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$, it follows that $f + g = g + f$.

(d) It is not necessary to always check for all the axioms, literally. Such *brute force* check works, but should only be taken to when there is no other way left. It is much better to use elegant reasoning, since not only it looks better, it takes lesser time to reproduce, is much more relevant as it casts light on your understanding, and is actually the essence of mathematics.

So, let us see how does one go about with the above problems using cute reasoning. You would be astonished to see that they take not more than a couple of lines to express what many of you wrote for pages — indeed I was delighted to see that there were some of you who were indeed *cute* in this respect. At this point, let me tell you, that I have penalised for not being cute, just to make you become better at *being* cute!

(a): $\mathbb{R}$ is a vector space with its usual operations of addition as vector addition and multiplication as scalar multiplication.

The operations on $\mathbb{R}^n$ is defined coordinatewise using the addition and multiplication of real numbers. Hence, they satisfy all the axioms and $\mathbb{R}^n$ is a vector space.

(b): Since $\mathbb{R}$ is a vector space and the operations in $\mathbb{R}^X$ are defined using the addition and multiplication of reals at each $x \in X$, the axioms are satisfied, so that $\mathbb{R}^X$ is a vector space.

the rest. . . : The rest of the problems depended on the simple observation: if V is a vector space, $\varnothing \subset S \subseteq V$ is a non-empty subset of V, $u, v \in S \Rightarrow \lambda u + \mu v \in S$, for every $\lambda, \mu \in \mathbb{R}$, then S is indeed a vector space.

In other words, I was using these problems to motivate the notion of a vector subspace which is actually specified later.

Here $C_1[0,1] \subseteq C[0,1] \subseteq \mathbb{R}^{[0,1]}$, $\mathbb{R}[x]_{\deg \le n} \subseteq \mathbb{R}[x] \subseteq C_1(\mathbb{R}) \subseteq C(\mathbb{R}) \subseteq \mathbb{R}^{\mathbb{R}}$, . . . , and you should now be able to complete the rest, . . .

PROBLEM 2. It is always important, in studying a mathematical structure, to be able to recognise an example as a proper one, and to recognise a pretender as one that fails in some respects. In each of the following, state with reasons which one is an example and which is a pretender.

(a) The set $\mathbb{R}$ of real numbers with usual addition of real numbers as vector addition and scalar multiplication defined by $(\lambda, x) \mapsto \lambda^2 x$. [1]
(b) The set $\mathbb{R}^2$ with vector addition being usual coordinate-wise addition and the scalar multiplication being defined by $(\lambda, (x, y)) \mapsto (\lambda x, 0)$.
(c) The set $\mathbb{Q}$ of rational numbers with usual addition of rational numbers as vector addition and the scalar multiplication being given by $(\lambda, x) \mapsto \lambda x$.
(d) The set $\mathbb{R}_{>0}$ of positive real numbers with usual multiplication of positive reals as the vector addition and the scalar multiplication being given by $(\lambda, x) \mapsto x^{\lambda}$.

[$4 \times 10 = 40$ marks]

SKETCHES OF SOLUTION 2. I was expecting the reasons only.

In (a) the scalar multiplication does not distribute over scalar addition, in (b) the unit law fails, in (c) the scalar multiplication is not defined. Only (d) is a vector space, and is due to the fact that multiplication of real numbers is commutative, associative, has 1 as identity and every non-zero real number has an inverse, and the properties of the scalar multiplication is exactly the laws of indices.

[1] This means that the scalar multiplication $\mathbb{R} \times V \xrightarrow{s} V$ on V is given by $s(\lambda, x) = \lambda^2 x$.

2. Video Lecture 2

PROBLEM 3. Which of the following are linear transformations and which are pretenders? Give reasons for your answers.

(a) $\mathbb{R} \xrightarrow{f} \mathbb{R}$ where $f(x) = 3x$.
(b) $\mathbb{R}^2 \xrightarrow{f} \mathbb{R}$ where $f(x, y) = x^2 + y^2$.
(c) $\mathbb{R}^2 \xrightarrow{f} \mathbb{R}^3$ where $f(x, y) = (3x + 4y, 7y - 4x, 2x)$.
(d) $\mathbb{R} \xrightarrow{f} \mathbb{R}$ where $f(x) = 3x + 7$.
(e) $[0, 1] \xrightarrow{f} \mathbb{R}$ where $f(x) = 2x$.
(f) $V \times W \xrightarrow{p_1} V$, where $p_1(v, w) = v$ with vector spaces V and W and the set $V \times W$ made into a vector space with vector addition defined by $(v, w) + (v', w') = (v + v', w + w')$ and scalar multiplication defined by $\lambda(v, w) = (\lambda v, \lambda w)$.

[$5 \times 6 = 30$ marks]

SKETCHES OF SOLUTION 3. Recall the definition of a linear transformation: a function $V \xrightarrow{f} W$ from a vector space V to a vector space W is a linear transformation, if it satisfies two conditions, namely:

$$f(x + y) = f(x) + f(y), \text{ for all } x, y \in V$$

and

$$f(\lambda x) = \lambda f(x), \text{ for all } \lambda \in \mathbb{R} \text{ and } x \in V.$$

The two conditions can be compressed into one equivalent condition, namely:

$$f(x + \mu y) = f(x) + \mu f(y), \quad \text{for all } x, y \in V \text{ and } \mu \in \mathbb{R}.$$

However, except for these two explicitly stated conditions there is also a very tacit condition for the function f to satisfy — the **domain** and the **codomain** of the function f **must** be both **vector spaces**.

Thus, (e) falls short of this inexplicit condition — it is not specified nor well known how can one consider $[0, 1]$ to be a vector space, and hence is **not** an example of a linear transformation.

If $V \xrightarrow{f} W$ is a linear transformation then $f(0) = 0$ — $f(0) = f(0 + 0) = f(0) + f(0) \Rightarrow f(0) = 0$. Since in (d) $f(0) = 7$ it is **not** an example of a linear transformation.

The others are as can be well verified. For instance, in the case of (f) the verification would run as follows:

given $(v,w),(v',w') \in V \times W$, $\lambda \in \mathbb{R}$:

$$p_1\big((v,w)+\lambda(v',w')\big) = p_1\big(v+\lambda v', w+\lambda w'\big) = v+\lambda v' = p_1(v,w)+\lambda p_1(v',w'),$$

implying the linearity of p_1.

PROBLEM 4. Let V and W be vector spaces and $\hom(V,W)$ be the set of all linear transformations from V to W, i.e.

$$\hom(V,W) = \left\{f : \text{the function } V \xrightarrow{f} W \text{ is a linear transformation}\right\}.$$

Define for $f,g \in \hom(V,W)$, $a \in V$ and $\lambda \in \mathbb{R}$:

$$(f+g)(v) = f(v)+g(v), \tag{1}$$

$$(\lambda f)(v) = \lambda f(v), \tag{2}$$

and

$$\hom(V,W) \xrightarrow{E_a} W \text{ , where } E_a(f) = f(a). \tag{3}$$

(a) Show that $\hom(V,W)$ with the vector addition defined in (1) and scalar multiplication defined in (2) is a vector space.

(b) Show also that the function $\hom(V,W) \xrightarrow{E_a} W$ from the vector space $\hom(V,W)$ to the vector space W is a linear transformation.

Taking the special case of $V = \mathbb{R} = W$, show that the function $\hom(\mathbb{R},\mathbb{R}) \times \mathbb{R} \xrightarrow{E} \mathbb{R}$ defined by $E(f,a) = f(a)$ is however **not** a linear transformation. Note that, using the result in Problem 3((f)) $\hom(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ above is indeed a vector space.

[$7+8+5 = 20$ marks]

SKETCHES OF SOLUTION 4. (a) This is very much similar to the proof of Problem 1(b), where you used $\mathbb{R}$, and here you need to use W.

More precisely, if W be a vector space then for any set X the set W^X of all functions $X \xrightarrow{f} W$ is a vector space with pointwise definitions of addition and scalar multiplication, as in (1) & (2).

It is enough to show that given the linear transformations $V \underset{g}{\overset{f}{\rightrightarrows}} W$ and a scalar $\lambda \in \mathbb{R}$ the function $h = f + \lambda g$, defined pointwise is also a linear transformation. Given this, it would follow that $\hom(V,W) \subseteq W^X$ (where $X = V$) with addition and scalar multiplication exactly the same as in W^V and hence is a vector space.

To show h is a linear transformation, consider $x,y \in V$ and $\mu \in \mathbb{R}$ and then:

$$\begin{aligned} h(x+\mu y) &= (f+\lambda g)(x+\mu y) \\ &= f(x+\mu y)+\lambda g(x+\mu y) \quad (\text{ using } 1\&2) \\ &= f(x)+\mu f(y)+\lambda(g(x)+\mu g(y)) \quad (\text{using linearity of } f\&g) \end{aligned}$$

$$= \big(f(x) + \lambda g(x)\big) + \mu\big(f(y) + \lambda g(y)\big) \quad \text{(rearranging the terms)}$$
$$= h(x) + \mu h(y) \quad \text{(using 1 \& 2)}$$

proves h to be linear.

(b) The function $\hom(V,W) \xrightarrow{E_a} W$ is defined by $E_a(f) = f(a)$, and could be called the *evaluation* of the linear transformation f at $a \in V$.

Take $f, g \in \hom(V,W)$ and any $\lambda \in \mathbb{R}$; to show $E_a(f + \lambda g) = E_a(f) + \lambda E_a(g)$ is just the usage of 1 & 2.

Note that the vector space $\hom(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ has operations: $(f,x) + \lambda(g,y) = (f + \lambda g, x + \lambda y)$;. . . the rest of finding the counter example should be simple.

Indeed, assuming that you have satisfied yourself about $\hom(\mathbb{R},\mathbb{R}) \times \mathbb{R} \xrightarrow{E} \mathbb{R}$ being not a linear transformation, one could go further to check that indeed it is bilinear — keep one coordinate fixed, and then it is linear in the other variable. Also, it is clear to check that $\dim \hom(\mathbb{R},\mathbb{R}) = 1$, so that what are the matrices for the linear transformations $\mathbb{R} \xrightarrow{E(f,-)} \mathbb{R}$ and $\hom(\mathbb{R},\mathbb{R}) \xrightarrow{E(-,x)} \mathbb{R}$, given any $f \in \hom(\mathbb{R},\mathbb{R})$ and $x \in \mathbb{R}$?

3. Video Lecture 3

PROBLEM 5. (a) Show that $S \subseteq W$ is a vector subspace of the vector space W, if and only if, for all $\lambda, \mu \in \mathbb{R}$ and every $s, t \in S$, $\lambda s + \mu t \in S$.

(b) Given the vector space $\mathbb{R}^3$ of triples of real numbers with the usual addition and scalar multiplication, which of the following subsets make a vector subspace:
(i) $S = \{(0, y, z) : y, z \in \mathbb{R}\}$.
(ii) $S = \{(x, y, z) : x = 0 \text{ or } y = 0\}$.
(iii) $S = \{(x, y, z) : x + y = 0\}$.
(iv) $S = \{(x, y, z) : x + y = 1\}$.

(c) Given the vector space $\mathbb{R}[x]$ of all polynomial functions with their usual addition and scalar multiplication which of the following subsets make a vector subspace of $\mathbb{R}[x]$.
(i) $S = \{p(x) : \deg p(x) = 3\}$.
(ii) $S = \{p(x) : 2p(0) = p(1)\}$.
(iii) $S = \{p(x) : 0 \leq x \leq 1 \Rightarrow p(x) \geq 0\}$.
(iv) $S = \{p(x) : p(x) = p(1 - x)\}$.

In each of the cases also evaluate $\operatorname{span}[S]$.

[$10 + 8 \times (2 + 3) = 50$ marks]

SKETCHES OF SOLUTION 5. (a) If $S \subseteq V$ is a vector subspace of V then it is trivial.

For the converse, observe that taking special choices of $s, t \in S$ and $\lambda, \mu \in \mathbb{R}$ the two conditions follow.

This is just a comprehension of the two conditions into one and a very simple checking rule.

(b) Every vector subspace must contain the zero vector. Hence (iv) is not a vector subspace.

For (ii), since $(2, 0, 3) + (0, 2, 3) = (2, 2, 6)$, the vectors on the left hand side are both in S and the one right is not in S it follows that S is not a vector subspace.

The rest are vector subspaces.

Regarding the span, let me illustrate with one, say the one in (iv). Clearly, in this case, $(x, y, z) \in S \Leftrightarrow x + y = 1$ implies the vector is of the form $(x, 1 - x, z)$ where $x, z \in \mathbb{R}$. Hence:

$$(x, 1 - x, z) = (0, 1, 0) + (x, -x, 0) + (0, 0, z) = (0, 1, 0) + x(1, -1, 0) + z(0, 0, 1),$$

implies that the vectors of S are precisely of the form:

$$S = \{(0, 1, 0) + x(1, -1, 0) + z(0, 0, 1) : x, z \in \mathbb{R}\}.$$

In other words, they are *translates* of vectors from the two dimensional subspace $\text{span}[(1, -1, 0), (0, 0, 1)]$ of $\mathbb{R}^3$ by the vector $(0, 1, 0)$.

Hence, a vector from the $\text{span}[S]$ would be a sum of two vectors, one being a multiple of $(0, 1, 0)$ and other being a vector from $\text{span}[(1, -1, 0), (0, 0, 1)]$. Hence $\text{span}[S] = \text{span}[(0, 1, 0), (1, -1, 0), (0, 0, 1)]$.

Since $\{(0, 1, 0), (1, -1, 0), (0, 0, 1)\}$ is a linearly independent set, and $\dim \mathbb{R}^3 = 3$, it follows that $\text{span}[S] = \mathbb{R}^3$.

(c) Since vector subspaces must:

- contain the zero vector, (i) is not a vector subspace,
- contain the negative of a vector, (iii) cannot be a vector subspace.

The rest should be simple.

Problem 6. (a) Let $S = \{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 1)\} \subseteq \mathbb{R}^3$.

Show that S is a linearly dependent set, but any set of three vectors from S is linearly independent.

What can you say about two element subsets or one element subsets of S — are they linearly independent or linearly dependent subsets?

(b) True or false: if $S = \{x, y, z\}$ be a linearly independent set of vectors in a vector space W then so also is the set $T = \{x + y, y + z, z + x\}$?

If it is true then prove it, else give an example to disprove it.

(c) Show that a set of two vectors (x_1, x_2, x_3) and (y_1, y_2, y_3) from $\mathbb{R}^3$ is linearly dependent, if and only if, the coordinates are in a constant ratio.

Can you provide a simpler condition (as above, in terms of the coordinates only) which equivalently states that three vectors (x_1, x_2, x_3), (y_1, y_2, y_3) and (z_1, z_2, z_3) of $\mathbb{R}^3$ are linearly dependent?

What can you say about a set of four vectors or more from $\mathbb{R}^3$?

(d) Find two bases B and C of $\mathbb{R}^4$ such that no vector is common to both B and C, while B contains the vectors $(1,0,0,0)$ and $(1,1,0,0)$ while C contains $(1,1,1,0)$ and $(1,1,1,1)$.

[10 + 10 + 10 + 10 = 40 marks]

Sketches of Solution 6. (a) Since $\dim \mathbb{R}^3 = 3$ and S has four vectors, it must be a linearly dependent set. *You need not do any calculations at all*!

There are four subsets of three vectors, namely:

$$S_1 = \{(1,0,0),(0,1,0),(0,0,1)\}, \qquad S_2 = \{(1,0,0),(0,1,0),(1,1,1)\},$$
$$S_3 = \{(1,0,0),(1,1,1),(0,0,1)\}, \qquad S_4 = \{(1,1,1),(0,1,0),(0,0,1)\}.$$

S_1 is a linearly independent set is immediate; the pattern suggests that proving one of the rest to be linearly independent would be enough for the others would be similar.

To show S_2 is linearly independent, use S_1:

$$\alpha(1,0,0) + \beta(0,1,0) + \gamma(1,1,1) = (0,0,0)$$
$$\Rightarrow (0,0,0) = \alpha(1,0,0) + \beta(0,1,0) + \gamma\big((1,0,0) + (0,1,0) + (0,0,1)\big) \dots$$
$$\Rightarrow \alpha + \gamma = \beta + \gamma = \gamma = 0$$
$$\Rightarrow \alpha = \beta = \gamma = 0,$$

proving the linear independence.

Every subset of a linearly independent set of vectors is linearly independent. Hence the rest follows, without any calculation!

Note: if you took to the computation route, then you would have to compute for $\binom{4}{3} = 4$ three element sets (which you reduced to two), $\binom{4}{2} = 6$ two element subsets and $\binom{4}{1} = 4$ one element subsets. Thus, observation and reasoning helped in reducing the number of computations from 14 to just 2.

(b) This was just used in the previous computation; now prove it in general.

(c) $(x_1, x_2, x_3) + \lambda(y_1, y_2, y_3) = (0,0,0) \Leftrightarrow \lambda = -\frac{x_1}{y_1} = -\frac{x_2}{y_2} = -\frac{x_3}{y_3}$ (if we could write it) and hence the conclusion.

Reason it out in full. . . .

Use this as a clue. . . .

PROBLEM 7. Let $V \xrightarrow{f} W$ be a linear transformation from the vector space V to the vector space W.

(a) Show that the set $\operatorname{Ker} f = \{v \in V : f(v) = 0\}$ is a vector subspace of V.[2]

(b) Show that the set $f[V] = \{w \in W : w = f(v) \text{ for some } v \in V\}$ is a vector subspace of W.[3]

[5 + 5 = 10 marks]

SKETCHES OF SOLUTION 7. (a) If $f(v) = 0 = f(u)$ and $\lambda \in \mathbb{R}$ then $f(u + \lambda v) = f(u) + \lambda f(v) = \ldots$

(b) Similar....

[2] $\operatorname{Ker} f$ is said to be the *kernel* of f.
[3] $f[V]$ is said to be the *image* of f.

4. Video Lecture 4

Problem 8. (a) Show that there exists a unique linear transformation $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}[x]_{\deg\le 5}$ such that:

$$f(1,0,1) = 1 + x^2, f(1,1,0) = 1 + x \text{ and } f(0,1,1) = x^3 + x^4.$$

Find $\operatorname{Ker} f$, $\mathrm{I}m[f]$ and their dimensions.

(b) Show that there exists a unique linear transformation $\mathbb{R}[x]_{\deg\le 2} \xrightarrow{f} \mathbb{R}^2$ such that:

$$f(1 - x + x^2) = (0,2) = f(2x^2) \text{ and } f(1 + x + x^2) = (2,2).$$

Find $\operatorname{Ker} f$, $\mathrm{I}m[f]$ and their dimensions.

(c) Does there exist a linear transformation $\mathbb{R}^3 \xrightarrow{f} \mathbb{R}^3$ such that:

$$f(1,0,1) = (1,2,1), f(0,1,1) = (2,1,1) \text{ and } f(-1,-1,-2) = (1,1,2)?$$

If it does, what is the formula for $f(x_1,x_2,x_3)$? what is $\operatorname{Ker} f$ and $\mathrm{I}m[f]$?

If it does not, then explain why does it not?

[5 + 5 + 5 = 15 marks]

Sketches of Solution 8. (a) Using the principle in Problem 6(b) (see page 10 for solution) it follows that $B = \langle (1,0,1),(1,1,0),(0,1,1)\rangle$ is a basis of $\mathbb{R}^3$.

Since:

$$(x,y,z) = \frac{x-y+z}{2}(1,0,1) + \frac{x+y-z}{2}(1,1,0) + \frac{-x+y+z}{2}(0,1,1),$$

it follows that $\operatorname{coord}_B((x,y,z)) = (\frac{x-y+z}{2}, \frac{x+y-z}{2}, \frac{-x+y+z}{2})$.

Hence the principle of linear extension tells us:

$$\begin{aligned} f(a,b,c) &= \begin{pmatrix} \frac{a-b+c}{2} & \frac{a+b-c}{2} & \frac{-a+b+c}{2} \end{pmatrix} \begin{pmatrix} 1+x^2 \\ 1+x \\ x^3+x^4 \end{pmatrix} \\ &= \frac{a-b+c}{2}(1+x^2) + \frac{a+b-c}{2}(1+x) + \frac{-a+b+c}{2}(x^3+x^4) \\ &= (a+b+c) + \frac{a+b-c}{2}x + \frac{a-b+c}{2}x^2 + \frac{-a+b+c}{2}x^3 + \frac{-a+b+c}{2}x^4. \end{aligned}$$

This describes the $\mathrm{Im}[f]$.

For the kernel, $f(a,b,c) = 0 \Leftrightarrow a-b+c = a+b-c = -a+b+c = 0 \Leftrightarrow a = b = c = 0$, so that $\operatorname{Ker} f = \{0\}$.

Hence $\dim \operatorname{Ker} f = 0$ and hence $\dim \operatorname{Im}[f] = \dim \mathbb{R}^3 = 3$.

(b) Similar...

(c) Since:

$$(1,0,1) + (0,1,1) + (-1,-1,-2) = (0,0,0)$$

the set $S = \{(1,0,1),(0,1,1),(-1,-1,-2)\}$ is a linearly dependent set.

If there were a linear transformation as suggested then $f(1,0,1)+f(0,1,1)+f(-1,-1,-2) = (0,0,0)$, which is not the case; hence there does not exist any linear transformation with the specified assignments.

PROBLEM 9. (a) Given the linear transformation $\mathbb{R}[x]_{\deg\le 2} \xrightarrow{f} \mathbb{R}^2$ in Problem 8.(b), find the matrix presentation $\operatorname{Mat}_{B\to C}(f)$ of f, given the basis $B = \langle 1-x, 1+x, x^2\rangle$ of $\mathbb{R}[x]_{\deg\le 2}$ and the basis $C = \langle (1,2),(2,1)\rangle$ of $\mathbb{R}^2$.

(b) It is easy to see that $B' = \langle 1, x, x^2\rangle$ and $C' = \langle (1,0),(0,1)\rangle$ are bases of $\mathbb{R}[x]_{\deg\le 2}$ and $\mathbb{R}^2$, respectively, and are possibly much more simpler.

Can you use change of basis matrices to give an elegant solution to the question in (a)?

[15 + 15 = 30 marks]

<u>SKETCHES OF SOLUTION</u> 9. (a) Just mere computation..., complete it...

(b) Surely:

$$\boldsymbol{P}_{B\to B'} = \begin{pmatrix}\operatorname{coord}_{B'}(1-x) & \operatorname{coord}_{B'}(1+x) & \operatorname{coord}_{B'}(x^2)\end{pmatrix} = \begin{pmatrix}1 & 1 & 0\\ -1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

and

$$\boldsymbol{P}_{C\to C'} = \begin{pmatrix}\operatorname{coord}_{(1,2)}(C') & \boldsymbol{P}_{C'\to 2,1}\end{pmatrix} = \begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}.$$

Since $\boldsymbol{P}_{B\to B'} = \operatorname{Mat}_{B\to B'}(\mathbf{1}_{\mathbb{R}[x]_{\deg\le 2}})$, $\boldsymbol{P}_{C\to C'} = \operatorname{Mat}_{C\to C'}(\mathbf{1}_{\mathbb{R}^2})$, and the required matrix is $\operatorname{Mat}_{B\to C}(f)$, it follows that:

$$\begin{aligned}
\operatorname{Mat}_{B\to C}(f) &= \operatorname{Mat}_{C'\to C}(\mathbf{1}_{\mathbb{R}^2})\operatorname{Mat}_{B'\to C'}(f)\operatorname{Mat}_{B\to B'}(\mathbf{1}_{\mathbb{R}[x]_{\deg\le 2}})\\
&= \begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}^{-1}\begin{pmatrix}\operatorname{coord}_{C'}(f(1)) & \operatorname{coord}_{C'}(f(x)) & \operatorname{coord}_{C'}(f(x^2))\end{pmatrix}\begin{pmatrix}1 & 1 & 0\\ -1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\\
&= \begin{pmatrix}-\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & -\frac{1}{3}\end{pmatrix}\begin{pmatrix}\operatorname{coord}_{C'}((1,2)) & \operatorname{coord}_{C'}((1,0)) & \operatorname{coord}_{C'}((0,1))\end{pmatrix}\begin{pmatrix}1 & 1 & 0\\ -1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}
\end{aligned}$$

$$= \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \quad \text{a } 2 \times 3 \text{ matrix, } \ldots.$$

PROBLEM 10. (a) Given the linear transformations $V \underset{g}{\overset{f}{\rightrightarrows}} W \xrightarrow{h} P$ between finite dimensional vector spaces, where $\dim V = m$, $\dim W = n$ and $\dim P = p$, show that for any given bases B, C, D of the vector spaces V, W, P respectively, and for any $\lambda \in \mathbb{R}$:

$$\text{Mat}_{B \to C}(f + g) = \text{Mat}_{B \to C}(f) + \text{Mat}_{B \to C}(g) \tag{4}$$

$$\text{Mat}_{B \to C}(\lambda f) = \lambda \text{Mat}_{B \to C}(f) \tag{5}$$

and

$$\text{Mat}_{B \to D}(h \circ f) = \text{Mat}_{C \to D}(h) \text{Mat}_{B \to C}(f). \tag{6}$$

<u>Hint:</u>
Recall from Problem 4, equations (1) & (2) on page 6:

$$(f + g)(v) = f(v) + g(v),$$

and

$$(\lambda f)(v) = \lambda f(v).$$

(b) Show that the set $\mathcal{GL}_{m,n}(\mathbb{R})$ of all $m \times n$ matrices with the usual matrix addition and scalar multiplication is a finite dimensional vector space with $\dim \mathcal{GL}_{m,n}(\mathbb{R}) = mn$.

Show that if V and W be finite dimensional vector spaces with $\dim V = m$ and $\dim W = n$, B a basis for V and C a basis for W then $\hom(V, W) \xrightarrow{\text{Mat}_{B \to C}(-)} \mathcal{GL}_{m,n}(\mathbb{R})$ is a bijective linear transformation.

Hence, or otherwise, obtain $\dim \hom(V, W)$.

Observe, in this context, something very interesting: there is a *multiplication*, $\mathcal{GL}_{p,n}(\mathbb{R}) \times \mathcal{GL}_{n,m}(\mathbb{R}) \to \mathcal{GL}_{p,m}(\mathbb{R})$ which is the usual matrix multiplication $(A, B) \mapsto AB$.

Alongside there is the *composition* of linear transformations $\hom(W, P) \times \hom(V, W) \xrightarrow{\circ} \hom(V, P)$ given by $(h, f) \mapsto h \circ f$.

None of these maps — neither the multiplication of matrices nor the composition of linear transformations, are linear transformations: say for matrices, it is known that $(A + C)(B + D) \neq AC + BD$ (check: this equation is exactly the demand for linearity of matrix multiplication). However, (6) on page 14 provides a connection between the bijective linear transformations $\mathrm{Mat}_{B\to D}(-)$, $\mathrm{Mat}_{C\to D}(-)$ and $\mathrm{Mat}_{B\to C}(-)$.

In the special case when $V = W = P$ and thence $m = n = p$, one has the matrix multiplication to be a binary operation on the vector space $\mathcal{GL}_{n,n}(\mathbb{R})$ (which is shortened to $\mathcal{GL}_n(\mathbb{R})$). This binary operation distributes over addition and scalar multiplication:

$$A(B + C) = AB + AC \qquad (B + C)A = BA + CA \tag{7}$$

and

$$A(\lambda B) = \lambda(AB) \qquad (\lambda A)B = \lambda(AB). \tag{8}$$

A vector space with an associative multiplication which distributes over addition and scalar multiplication as in (7) & (8) and which also possesses an identity is often called an *real unital algebra*, and $\mathcal{GL}_n(\mathbb{R})$ as well as $\hom(V, V)$ are examples of such. Furthermore, $\mathrm{Mat}_{B\to B}(-)$ then shows that $\hom(V, V)$ and $\mathcal{GL}_n(\mathbb{R})$ are *same* as real unital algebras. However, we shall not be discussing algebras any further, it just came up incidentally with this problem.

Furthermore, the operations of matrix multiplication or composition of linear transformations are *not very far* from being linear transformations. For instance, each of the following maps:

(i) for each fixed matrix A, the map $B \mapsto AB$,
(ii) for each fixed matrix B, the map $A \mapsto AB$,
(iii) for each fixed linear transformation h, the map $f \mapsto h \circ f$,
(iv) for each fixed linear transformation f, the map $h \mapsto h \circ f$,

are indeed linear transformations (check it...). Such maps are often called *bilinear maps* and we shall come back to them later on.

If you continue with your studies with mathematics, you shall soon learn that *bilinear* maps correspond to linear transformations, but **not** from the product vector space, but from a vector space very closely related to it, called the *tensor product* of the two vector spaces.

(c) Given the finite dimensional vector space V with $\dim V = n$ and the bases P, Q, R, show that:

(9) $$\boldsymbol{P}_{P\to R} = \boldsymbol{P}_{Q\to R}\boldsymbol{P}_{P\to Q}.$$

Hence, or otherwise, show that $\boldsymbol{P}_{P\to Q}$ is an invertible matrix and ${\boldsymbol{P}_{P\to Q}}^{-1} = \boldsymbol{P}_{Q\to P}$.

$[[5+5+5]+[5+3+2]+[8+2] = 35$ marks$]$

<u>SKETCHES OF SOLUTION</u> 10. (a) Clearly, if we take $B = \langle v_1, v_2, \ldots, v_n\rangle$, $C = \langle w_1, w_2, \ldots, w_m\rangle$ and $D = \langle u_1, u_2, \ldots, u_p\rangle$

$$\mathrm{Mat}_{B\to C}(f) = \begin{pmatrix}\mathrm{coord}_C(f(v_1)) & \mathrm{coord}_C(f(v_2)) & \ldots & \mathrm{coord}_C(f(v_n))\end{pmatrix},$$

$$\mathrm{Mat}_{B\to C}(g) = \begin{pmatrix}\mathrm{coord}_C(g(v_1)) & \mathrm{coord}_C(g(v_2)) & \ldots & \mathrm{coord}_C(g(v_n))\end{pmatrix},$$

and

$$\mathrm{Mat}_{C\to D}(h) = \begin{pmatrix}\mathrm{coord}_D(h(w_1)) & \mathrm{coord}_D(h(w_2)) & \ldots & \mathrm{coord}_D(h(w_m))\end{pmatrix}.$$

Since $(f+g)(v_i) = f(v_i) + g(v_i)$ and $(\lambda f)(v_i) = \lambda f(v_i)$, for $i = 1, 2, \ldots, n$:

$$\begin{aligned}\mathrm{Mat}_{B\to C}(f+g) &= \begin{pmatrix}\mathrm{coord}_C((f+g)(v_1)) & \mathrm{coord}_C((f+g)(v_2)) & \ldots & \mathrm{coord}_C((f+g)(v_n))\end{pmatrix}\\ &= \mathrm{Mat}_{B\to C}(f) + \mathrm{Mat}_{B\to C}(g),\end{aligned}$$

and

$$\begin{aligned}\mathrm{Mat}_{B\to C}(\lambda f) &= \begin{pmatrix}\mathrm{coord}_C((\lambda f)(v_1)) & \mathrm{coord}_C((\lambda f)(v_2)) & \ldots & \mathrm{coord}_C((\lambda f)(v_n))\end{pmatrix}\\ &= \lambda \mathrm{Mat}_{B\to C}(f).\end{aligned}$$

If $\mathrm{coord}_C(f(v_i)) = (a_1, a_2, \ldots, a_m)$ then $h(f(v_i)) = a_1 h(w_1) + a_2 h(w_2) + \ldots a_m h(w_m)$, for each $i = 1, 2, \ldots m$. Hence:

$$\mathrm{Mat}_{B\to D}(h \circ f) = \mathrm{Mat}_{C\to D}(h)\mathrm{Mat}_{B\to C}(f).$$

(b) Let for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, A_{ij} be the $m \times n$ matrix whose i-jth element is 1 and the rest are all 0.

Show that:

$$B = \left\{A_{ij} : 1 \le i \le m \text{ and } 1 \le j \le n\right\}$$

is a basis for $\mathcal{GL}_{m,n}(\mathbb{R})$.

The fact that $\mathrm{Mat}_{B\to C}(-)$ is a linear transformation is attested by (4) & (5).

The coordinatisation principle actually states in other words that this is actually a bijection.

Since $\dim \mathcal{GL}_{m,n}(\mathbb{R}) = mn$, and bijective linear transformations preserve dimensions, $\dim \hom(V, W) = mn$.

(c) Simple...

5. (Video) Lecture 5

Problem 11 (Two Almost Missed, yet Essential Facts About Matrices). Given a matrix of order $m \times n$, say:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} = \left(a_{ij}\right)_{i,j=1}^{m,n},$$

its rows are $\mathbf{row}_i[A]$, for $i = 1, 2, \dots, m$ are determined by elements of $\mathbb{R}^n$ and columns are $\mathbf{col}_j[A]$, for $j = 1, 2, \dots, n$ are determined by elements of $\mathbb{R}^m$, where:

$$\mathbf{row}_i[A] = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{R}^n, i = 1, 2, \dots, m$$

and

$$\mathbf{col}_j[A] = (a_{1j}, a_{2j}, \dots, a_{mj}) \in \mathbb{R}^m, j = 1, 2, \dots, n.$$

As a convention we shall always denote an element of $\mathbb{R}^n$ by a $n \times 1$ matrix, i.e., as a column vector.

(a) Show that for any $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$:

$$\mathbf{y} = A\mathbf{x} = x_1\mathbf{col}_1[A] + x_2\mathbf{col}_2[A] + \dots + x_n\mathbf{col}_n[A] \in \operatorname{span}[\mathbf{col}_1[A], \mathbf{col}_2[A], \dots, \mathbf{col}_n[A]].$$

(b) Hence show that for any matrix $B = \left(b_{ij}\right)_{i,j=1}^{n,p}$ of order $n \times p$ the columns of AB are $A\mathbf{col}_1[B], A\mathbf{col}_2[B], \dots, A\mathbf{col}_p[B]$.

Problem 12 (Rank and Nullity for Matrices). Recall from **Video Lecture 4** the finite dimensional vector spaces are, under coordinatisation, precisely some $\mathbb{R}^n$ $(n \geq 0)$ and from Problem 10(b) (see page 14) that the linear transformations $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$ are precisely the $m \times n$ matrices $F = \left(\phi_{ij}\right)_{i,j=1}^{m,n}$ with real entries. The same problem also illustrates the near match between the matrix operations and operations on linear transformations.

Thus, the linear transformation corresponding to a matrix F of order $m \times n$ is $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m$ defined by $f(x_1, x_2, \ldots, x_n) = F \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, where the point $(x_1, x_2, \ldots, x_n)$, in matrix form, shall be represented by the $n \times 1$ column matrix $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

(a) Given the $m \times n$ matrix F, show that the kernel of the associated linear transformation is the vector subspace:

$$K_F = \left\{ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : F \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0} \right\}$$

The vector subspace K_F is:

(i) The set of all solutions of the homogeneous system of linear equations determined by the matrix F.

(ii) Often called the *null space* of F.

(iii) The dimension of K_F is often called the *nullity* of the matrix F.

(b) Given a $m \times n$ matrix show that the image of the associated linear transformation is:

$$R_F = \left\{ F \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ v_n \end{pmatrix} : (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \right\}.$$

The vector subspace R_F is:

(i) Often called the *range space* of F or also the *column space* of F (see Problem 11 for an explanation of this name).

(ii) The dimension of R_F is often called the *rank* of the matrix F.

(iii) Thus, we now know that the sum of the rank and nullity of a matrix is equal to the number of columns of the matrix.

Sketches of Solution 11. Simple, . . . , complete them yourself. . . .

PROBLEM 13 (Solutions of Linear Equations). Given $V \xrightarrow{f} W$ a linear transformation from the vector space V to the vector space W, a fixed vector $b \in W$, the equation:

$$f(x) = b, \tag{10}$$

and the set $S = \{x \in V : f(x) = b\}$ of solutions of (10).

(a) Show that $S \neq \varnothing$, if and only if, $b \in f[V]$, the image of f.

(b) If $x_0 \in S$ then $x \in S$, if and only if, there exists a $t \in \operatorname{Ker} f$ such that $x = x_0 + t$.

> The special element $x_0 \in S$ is often called a *particular solution* of the equation in (10).
>
> Thus, the Moral is: solutions of linear equations, if they exist, are just a *translate* of a particular solution of the kernel of the linear transformation — observe the geometric language in use here, *translations* meaning as if the *kernel is pushed by the particular solution* thereby getting *translated* to the particular solution.

(c) Now suppose that both V and W are finite dimensional vector spaces with $\dim V = n$, $\dim W = m$; the linear transformation is then determined by a matrix F of order $m \times n$, and is given by $(x_1, x_2, \dots, x_n) \mapsto F \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.

The linear equation in (10) is then obtained on fixing an element $\mathbf{b} = (b_1, b_2, \dots, b_m) \in \mathbb{R}^m$ and trying to obtain the n-tuples $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ such that $F\mathbf{x} = \mathbf{b}$. The set of solutions is then $S = \{\mathbf{x} \in \mathbb{R}^n : F\mathbf{x} = \mathbf{b}\}$.

(i) Show that $S \neq \varnothing$, if and only if, the rank of the matrix F and the matrix $A = \begin{pmatrix} F & \mathbf{b} \end{pmatrix}$ are equal, where A is a matrix of order $m \times (n+1)$ with the first n columns as F and the last column as $\mathbf{b}$.

(ii) Show that if $\mathbf{x_0} \in S$ then $\mathbf{x} \in S$, if and only if, there exists a $\mathbf{t} \in \mathbb{R}^n$ such that $F\mathbf{t} = \mathbf{0}$ and $\mathbf{x} = \mathbf{x_0} + \mathbf{t}$.

> Note how this provides a computational justification to the process of solving linear systems of equations.

SKETCHES OF SOLUTION 12. Simple, ..., complete them yourself....

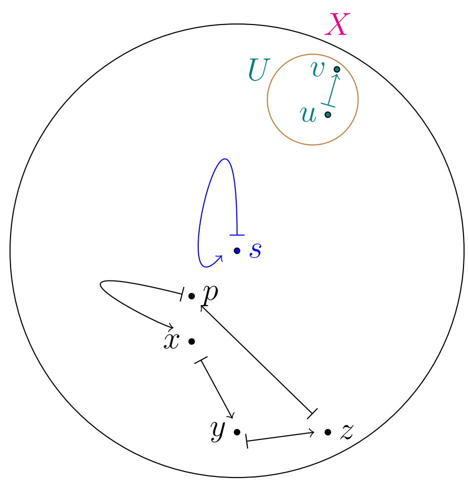
6. (Video) Lecture 6

This (video) lecture conveys the following:

(a) First starts with the idea of an *invariant subspace*, an idea which is of paramount importance in the study of linear algebra.

To say it in brief: given a function $X \xrightarrow{f} X$ on a set X (often such functions are called *colorblue endomaps* or *endomorphisms*) it is interesting to find a *fixed point* of such a function, which is an element $x \in X$, such that $f(x) = x$.

Every endomorphism on a set X ushers in a *dynamical picture*, a notion of a *motion*, in which we say that f *shifts* $x \in X$ to $y \in X$, if $y = f(x)$. This suggests the following diagram:



wherein you could think of possible *orbits* like $x \overset{f}{\longmapsto} y \overset{f}{\longmapsto} z \overset{f}{\longmapsto} p \overset{f}{\longmapsto} x$ which might be *finite*, or which could carry on and on..., and also *loops* like $s \overset{f}{\longmapsto} s$, which are exactly the *fixed points*.

You could also have a subset $U \subseteq X$ such that: for all $u \in U$, $v = f(u) \in U$; in other words, $f[U] \subseteq U$. Thus, once if we land up in such a set as U, we never *come out* of it — such sets U are obviously called *invariant subsets*.

The situation in vector spaces is similar, but we consider linear transformations $X \xrightarrow{f} X$ instead, and then we have invariant vector subspaces.

Every endomorphism of a vector space has at least the two trivial subspaces as its invariant subspaces. However, not every endomorphism of vector spaces have a non-trivial invariant subspace. Starting with the existence of an invariant subspace of dimension one of a finite dimensional vector space leads to the notion of an eigenvalue, eigenvector and eigenspaces. The eigenspaces turn out to be invariant subspaces.

(b) The notion of an eigenvalue of a square matrix (= linear transformation on a finite dimensional vector space) is closely connected to the characteristic polynomial.

Since every odd degree polynomial function in $\mathbb{R}[x]$ has a root, every odd order square matrix has an invariant subspace, while even order square matrices may or may not.

(c) The eigenvectors corresponding to distinct eigenvalues are linearly independent. This gives rise to the notion of *diagonalisation* of a square matric, which is then completely described in terms of its eigenvalues and eigenvectors.

(d) The notions of algebraic and geometric multiplicities are required in the complete formulation of diagonalisability, which are also discussed.

Problem 14. Two square matrices A and B are said to be *similar*, if there exists an invertible square matrix P such that $B = P^{-1}AP$.

(a) Interprete the relation of similarity of matrices in terms of *change of basis* and deduce that similar matrices represent the same linear transformation.

(b) Show that any two similar matrices A and B have the same characteristic polynomial, and hence the same eigenvalues.

(c) If $B = P^{-1}AP$ for some invertible square matrix P, show that for any eigenvalue λ of A, $\mathcal{E}_\lambda[B] \xrightarrow{P} \mathcal{E}_\lambda[A]$ sets up a bijective linear transformation between the eigenspaces.

[5 + 5 + 5 = 15 marks]

Sketches of Solution 13. You can now see that each matrix A of order $m \times n$ is actually representing a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ with respect to the standard basis on each side — recall that the standard basis for $\mathbb{R}^n$ is $U = \langle \mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n} \rangle$, where $\mathbf{e}_i \in \mathbb{R}^n$ has 0 in every coordinate except the ith one which is 1.

Hence, in this regard, each non-singular matrix A of order $n \times n$ is actually $\boldsymbol{P}_{B \to U}$, where $B = \langle \mathbf{col}_A[1], \mathbf{col}_A[2], \ldots, \mathbf{col}_A[n] \rangle$. Use this to interprete (a).

Regarding (b), the equation $\det(P^{-1}AP - \lambda I) = \det P^{-1}(A - \lambda I)P = \det P^{-1} \det(A - \lambda I) \det P = \det(A - \lambda I)$ should provide your answer.

Finally (c) is just a computation.

Problem 15. Let A be a square matrix of order n and $\lambda \in \mathbb{R}$ be an eigenvalue of A of geometric multiplicity k.

Choose a basis $B_0 = \langle \mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k} \rangle$ of $\mathcal{E}_\lambda[A]$ and extend this to a basis B of $\mathbb{R}^n$.

(a) Show that:

$$\mathrm{Mat}_{B\to B}(A) = \begin{pmatrix} \lambda I_k & P \\ \mathbf{0}_{n-k,k} & Q \end{pmatrix},$$

where P is a matrix of order $k \times (n-k)$, Q is a matrix of order $(n-k)\times(n-k)$ and $\mathbf{0}_{n-k,k}$ is the zero matrix of order $(n-k)\times k$.

(b) Show that the characteristic polynomial of $S = \begin{pmatrix} \lambda I_k & P \\ \mathbf{0}_{n-k,k} & Q \end{pmatrix}$ is:

$$\chi_S(x) = (x-\lambda)^k \chi_Q(x),$$

where for any square matrix T, $\chi_T \in \mathbb{R}[x]$ is the characteristic polynomial of the matrix T.

(c) Hence, otherwise, deduce that the algebraic multiplicity of λ is at least k its geometric multiplicity.

Hint: Use Problem 14.

[10 + 10 + 10 = 30 marks]

SKETCHES OF SOLUTION 14. Just use the definitions developed so far. . . .

The purpose of this is to show that: algebraic multiplicity of an eigenvalue is greater than or equal to its geometric multiplicity.

Note: Hence, if the algebraic multiplicity of an eigenvalue is 1 then you need not do any further computations regarding its geometric multilicity, since its eigenspace being non-trivial would have to be one dimensional and hence its geometric multiplicity would automatically be 1. You need to compute for only those eigenvalues which have algebraic multiplicity > 1.

PROBLEM 16. Given any polynomial function $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ $(a_n \neq 0)$ in $\mathbb{R}[x]$ and any square matrix A of order n, define:

$$f(A) = a_0I_n + a_1A + a_2A^2 + \cdots + a_nA^n.$$

Obviously, $f(A)$ is a square matrix of order n and hence a linear transformation on $\mathbb{R}^n$.

(a) If $f(x) = -1 + 2x + 3x^3$, evaluate $f(A)$, where:

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}.$$

(b) Show that for any $f \in \mathbb{R}[x]$, any eigenvalue $\lambda \in \mathbb{R}$ of a square matrix A and any $\mathbf{v} \in \mathcal{E}_\lambda[A]$:

$$f(A)\mathbf{v} = f(\lambda)\mathbf{v}.$$

(c) Hence, or otherwise, show that if A be a diagonalisable square matrix and $\chi_A \in \mathbb{R}[x]$ be its characteristic polynomial then $\chi_A(A) = \mathbf{0}_{n,n}$ the zero matrix of order $n \times n$, i.e., the zero linear transformation on $\mathbb{R}^n$.

Hint: Use the fact that a linear transformation $V \xrightarrow{f} W$ on a finite dimensional vector space V is identically zero, if and only if, for any basis B if V, $v \in B \Rightarrow f(v) = 0$. (Can you prove this using the Method of Linear Extension?)

(d) If A be a diagonalisable matrix, then compute $f(A)$ for any $f \in \mathbb{R}[x]$.
Can you give an algorithm to find A^{-1} in such a case?

[$5 + 5 + 10 + (5 + 10) = 35$ marks]

Sketches of Solution 15. Simple..., just use the definitions...and compute....

References

Department of Mathematical Sciences, Unisa Science Campus, corner of Christiaan de Wet & Pioneer Avenue, Florida 1709, Johannesburg, Gauteng, South Africa

E-mail address: ghoshpp@unisa.ac.za