I was looking for something in my old stack of books when I stumbled on to Sheldon Axler's fantastic book, "Linear Algebra Done Right". I have fond memories about the book. I think the last time I referred to this book was more than 3 years ago. Took a few hours to go over the book again. Like wine that tastes better when aged, I think some books also give the same kind of effect, atleast to me. A legitimate understanding of ANY discipline needs one to have a good grip on Linear Algebra. The more randomness in the field you choose to work with, the more you will use linear algebra, as you are forever looking for an approximate solution.

A very basic example that any stats person learns is least squares. When you build a model for $Y = X\beta + \epsilon$, all you are doing is to find the best $|| \cdot ||_2$ approximation of Y in the column space of X. Take any machine learning algorithm from the most basic to the most sophisticated, linear algebra pervades everywhere. Any Electrical engineer breathes FFT algo which is nothing but a change of basis. You take a discretized function in one basis and represent it in another basis. For a structural engineer, his/her job would certainly involve solving differential equation relating to structures and if they want to solve it numerically, they will need to study the stability of first order and second order difference matrices. An econometrician needs to understand invariant subspaces, i.e. the concept of eigen values and eigen vectors to study even the most elementary time series AR(1). Optimization pervades many fields and all the techniques involved in optimization are formulated, solved and reported in terms of linear algebra terminology. Minimizing a multivariate function entails computing gradients(vectors) and hessian(matrices).With out the language of Linear Algebra, optimization as a field would not have developed to what it is today. I can go on and on ...

But the point I want to make is this : For some reason, I had never learnt the importance of linear algebra as a "thinking tool" for many years in to my working life. When I did start learning the subject, I realized that it gives one a set of fresh eyes to look at a ton of stuff. Many things start to make sense when viewed from a matrix algebra perspective. For example, if you are asked to approximate the $f(x) = \sin x$ over $[-\pi, \pi]$ using a fifth order polynomial, $a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_tx^5$, how would you go about it ? What are your first thoughts ? Pause for a while ... Well, you can use Taylor series approximation but it only works near origin and not across the entire interval $[-\pi, \pi]$. A person who breathes stats would generate a few data points of $\sin x$ and do a polynomial regression. An electrical engineer would probably consider the 6 functions $(1, x, x^2, x^3, x^4, x^5)$ as input signals and then come up with a linear filter to approximate it to $\sin x$ function . Basically there are many ways to do it. However I think an elegant way to do is to think about the function $\sin x$ as infinite order polynomial and express the function in terms of the basis of a subspace, i.e. fifth order polynomial space. You can do this via pen and paper. No fancy software is needed to solve the problem. I like this book because this was the first book that really taught me what a matrix actually means. I will list down a few questions here. These questions drill home the importance of actually "understanding matrices" as oppposed to merely "using matrices". These questions are in no particular order of difficulty. One thing that is common across all of them is that Axler's book provides a beautiful explanation to all the questions and in the process will make anyone appreciate the beauty of Linear Algebra.

1. Why is matrix multiplication done the way it is ? Every one learns about matrix multiplication at some stage, school/ college/ undergrad etc. But only few ask the rationale behind such a procedure of rows multiplied columns in a matrix multiplication ? Why can't there be some other procedure to do matrix multiplication ? Why can't we do an element wise multiplication ? Well, Hadamard product is exactly that sort of element wise multiplication, however one reads/learns about such a kind of multiplication only in advanced courses, after the usual Cayley multiplication is learnt. Why can't one conjure up some

rule for multiplying matrices that allows multiplying matrices of different dimensions, like multiplying a 3 by 4 matrix with 1 by 2 matrix. All crazy things can be thought of relating, to matrix multipication. But we all learn one specific way to multiply matrices. What's the rationale ?

2. What is a linear map ?
3. What do you understand by a vector space ? What are the properties of a vector space ?
4. When is a vector space called finite dimensional ?
5. When are vectors called linearly independent ?
6. What is the difference between basis and a set of linearly independent vectors ?
7. Can every spanning list in a vector space be reduced to a basis of the vector space ?
8. What do you understand by the dimension of a vector space ?
9. There is a matrix behind every linear map. What does the statement mean ?
10. If there is a linear mapping from a subspace to itself, it is called an invariant space. Why should one be thinking about invariant spaces ?
11. What is the connection between invariant subspace of a linear map and eigen value of the linear map ?
12. What are called operators ?
13. What are eigen vectors ? What is the relation between eigen vector of an linear map and invariant subspace of a linear map ?
14. Every operator on a finite-dimensional, non zero vector space has an eigen value.Can you prove it ?
15. Why do we care so much about upper triangular matrices ? What is their connection with the matrix of a linear map ?
16. What do we achieve by finding a linear map that has a upper triangular matrix representation with respect to some basis ?
17. What do we achieve by finding a linear map that has a diagonal matrix representation with respect to some basis ?
18. What are inner product spaces ?
19. What is a norm ?
20. What does inner product inducing norm mean ?
21. Can every norm be induced by some inner product definition ?
22. Is normed space subset of inner product space or inner product space a subset of normed space ?
23. Given an example of a norm that is not induced by an inner product definition ?
24. What is the benefit of appending a vector space with an inner product definition ?
25. What are orthonormal basis ?
26. If V is a complex vector space and T is a linear mapping, then T must necessarily have an upper triangular matrix with respect to some orthonoral basis of V
27. What is a linear functional ?
28. What exactly is the rationale behind taking a transpose matrix ? What is the interpretation behind a transpose of a matrix ?
29. What are self-adjoint operators?
30. When is an operator on an inner product space called normal ?
31. A normal operator is norm preserving. What does it mean ?
32. For any normal operator on a space, there is atleast one eigen value. What does the statement mean ? Can you relate it to some application setting ?
33. V is a real inner product space and T is an operator that is adjoint. In other words, we are taking about symmetric real matrices. Why should there always be an orthonormal basis for this operator ?

34. What are positive semi definite matrices ? This question is not the right one to ask. The actual question should be, What are positive semi definite operators ? If you know this, then PSD matrices are just a representation of this linear map ?
35. Cholesky is one of the basic matrix factorization theorems that is taught in many applied courses. Why does cholesky work?
36. Square root of a matrix is often seen in books. It basically means that it is an matrix associated with an operator, called a square root operator. Can you name atleast a few areas where square root operators play a vital role ?
37. When is an operator supposed to have isometry ?
38. From a computational standpoint, what is the use if an operator has isometry ?
39. What are orthogonal operators ?
40. What are unitary operators ?
41. The book gives nice analogies between $\mathbb{C}$, the space of complex numbers and $\mathcal{L}(V)$. Some of them are :
    - A complex numer z corresponds to an operator T.
    - A complex conjugate $\overline{z}$ corresponds to an operator T^*.
    - The real numbers correspond to self-adjoint operators
    - The non negative numbers correspond to positive operators

    What do you think is an appropriate analogy for an operator being an isometry ?
42. The polar decomposition states that each operator on V is the product of an isometry and a positive operator. What is the analogy with the space of complex numbers ?
43. What do you mean by singular vaues of an operator ?
44. What does it mean when one says that an operator has a Singular value decomposition ?
45. Verbalize SVD without resorting to a lot of matrix jargon, i.e. use operator vocabulory to state SVD.
46. What are singular values ?
47. What is an hermitian operator ?
48. Is every normal matrix a self-adjoint matrix ?
49. Can the eigen value of an self-adjoint operator be complex?
50. If an operator is normal , what can you can say about the eigen vectors of the operator ?
51. What is a positive operator ?
52. Does a positive operator on a vector space have a unique positive square root ?
53. What is a generalized eigen vector of an operator?
54. Can the space of complex vector space be decomposed using eigen vectors ?
55. Can the space of complex vector space be decomposed using generalized eigen vectors ?
56. What is a nilpotent operator ?
57. What is the sum of multiplicities of an eigen value in a complex vector space ?
58. What is a characteristic polynomial of an operator?
59. What is a minimal polynomial of an operator ?
60. What do you understand by Jordan basis ? Why are they important ?
61. Jordan form is absolutely essential to understand theoretical aspects of Dynamic Linear Models. Can you take a guess why is that so ?
62. What do you intuitively understand by trace of a matrix ? Where could it find an application ?
63. What do you intuitively understand by determinant of a matrix ? Where could it be useful ?
64. Can you think of atleast 3 different methods to compute the determinant of a matrix ? Same with the trace ?