This week we will do two problem sessions since we did not get to see any section 2.2 problems. First, let me remark that the way I like to think about the matrix of a linear transform is I like to think of the linear transform $T$ as providing me instructions on how to map a vector. The matrix $[T]_{\beta}^\gamma$ can then be thought of as providing me the same instructions in a different language. (A language that depends on the choices of bases $\beta$ and $\gamma$.) Specifically, if I know the vector $v$ itself, then I could compute $Tv$ since I know $T$ and $v$. But if instead the information I have for $v$ is not what it is, but rather what its coordinates relative to $\beta$ are, then $[T]_{\beta}^\gamma$ tells me not directly what $Tv$ is but rather $[Tv]_{\gamma}$ is. I say “not directly” because you can still deduce what $Tv$ is, just as you can deduce what $v$ is. Knowledge of the coordinates of a vector tell you exactly what it is via the top of p.80 of your textbook, and knowledge of a vector tells you what the coordinates are by solving a system of linear equations. Because $\beta$ is a basis, and as is $\gamma$, you know that $v$ and $Tv$ respectively have unique coordinates with respect to $\beta$ and $\gamma$, so basically description of the actual vector is exactly as informative as describing its coordinates relative to a known, fixed basis.

Now, let’s look at problem 16 of section 2.2.

So you start playing with some ideas in your head. Here are the things that occurred to me first. (I had to think about this problem; it is not easy.) First I thought: let’s just try a “greedy” algorithm, which is where I make a selection of the vectors that will be in my eventual basis $\beta$ at each step, and I require $\gamma$ to consistent of the map $T$ applied to the choices I make for $\beta$, attempting to maintain that $\gamma$ needs to be linearly independent. Since I will make $\text{dim}(V)$ choices of $\beta$ and $\text{dim}(W) = \text{dim}(V)$, if I can keep $\gamma$ linearly independent as this algorithm proceeds, then I am done. However, it is not so easy. Basically, via this algorithm, although at each step you are trying to keep the independence, you could make a choice of vector for $\beta$ at some point that steps on your toes for the future. So you have to choose more judiciously than just going step by step.

Once I realized this, I thought: well, what do I know about how to get a matrix into diagonal form? Oh, reduced row echelon form. I should do row reductions, and show that all 3 row reduction operations (adding rows, rescaling them, switching them) preserve the property of having a choice of $\beta, \gamma$ that solve this problem. Then once its in diagonal form, it will be clear that such a choice $\beta, \gamma$ exists, being that the matrix is already diagonal.

Actually, this method doesn’t work either, because the reduced row echelon form may not be diagonal if the matrix has degeneracies. And it was at this point that I realized what had to be done. How do you account for degeneracies? You should be thinking of nullity now. So, much like the proof of Rank Nullity Theorem, we obtain a basis $\{v_1, \ldots, v_k\}$ for $N(T)$ where $k$ is the nullity of $T$. Extend this to a basis $\beta$ of $V$, where $V$ is the domain of $T : V \to W$. Then, if you review the proof of rank nullity, you see that they demonstrate that the image via $T$ of the vectors in $\beta$ that are not in $N(T)$ forms a basis for $R(T)$, which we will then extend to be a basis $\gamma$ of $W$. Come to section to see why this choice is acceptable.

Section 2.3 problem 11.
If $T^2 = T_0$ then that means that $\forall x \in V$ we have that $T(T(x)) = T^2(x) = T_0(x) = 0_V$ so that one learns that $T(x)$ is a vector which $T$ takes to 0, so that $T(x) \in N(T)$. Now, let $y \in R(T)$ so that $\exists x_0$ such that $y = T(x_0)$ and plug in $x_0$ into the statement we just derived. Then we learn that $y = T(x_0) \in N(T)$. This demonstrates that $R(T) \subset N(T)$ since we showed that any $y \in R(T)$ is also in $N(T)$.

On the other hand, suppose that $R(T) \subset N(T)$. Then $\forall x \in V$ we see that $T^2(x) = T(T(x))$. However, we already know that $T(x) \in R(T)$ by definition of the range of $T$. Thus, by our assumption, we learn that $T(x) \in N(T)$. This lets us know that $T(T(x)) = 0_V$, from which we learn that $T^2(x) = 0_V = T_0(x)$. Thus, since $x$ was arbitrary we learn that $T^2 = T_0$.

Section 2.4 Problem 4
To check that $AB$ has an inverse, we need to find some matrix $C$ for which $(AB)C = C(AB) = I$. Luckily, we don’t actually have to “find” it as they have given you the guess already. Let $C = B^{-1}A^{-1}$. Observe that the following computations are valid:

$$(AB)C = (AB)(B^{-1}A^{-1}) = A(B(B^{-1}A^{-1})) = A((BB^{-1})A^{-1}) = A(A^{-1}) = AA^{-1} = I$$

and

$$C(AB) = (B^{-1}A^{-1})(AB) = ((B^{-1}A^{-1})A)B = (B^{-1}(A^{-1}A))B = (B^{-1}(I))B = B^{-1}B = I.$$ 

In each of these, we have used the associative property of matrix multiplication twice.

Section 2.4 Problem 5:
Again, they have been kind enough to supply the guess, so the following computation gives one of the required, and I leave the other one to you.

$$(A^{-1})^t A^t = (A(A^{-1}))^t = I^t = I.$$ 

Here we have used how matrix multiplication and transposition interact. How do you do the other half?

Section 2.4 Problem 7:
This will be done in class only, and only if there’s time this week.