Linear Algebra – Notes on Interesting Topics

Arin Chang

Spring 2020

The following are a collection of interesting results and connections in Linear Algebra that I've compiled, mostly from Sheldon Axler's *Linear Algebra Done Right* as well as Ken Ribet's Spring 2020 lectures from Math 110 at UC Berkeley.

1 Riesz Representation Theorem

Theorem 1. *Suppose V is a finite dimensional inner product space and φ is a linear functional on V. Then there is a unique vector $u \in V$ such that $\varphi(v) = \langle v, u\rangle$ for every $v \in V$.*

Proof. To prove existence, take $e_1, ..., e_m$ to be an orthonormal basis of V. We can get an othonormal basis of V by constructing a regular basis of V and applying the Gram Schmidt process to turn it into an orthonormal one. Take $v = a_1e_1 + ... + a_me_m = \sum_{i=1}^{m}\langle v, e_i\rangle e_i$. Then apply φ to both sides:

$$\begin{aligned} \varphi(v) &= \varphi(\sum_{i=1}^{m}\langle v, e_i\rangle e_1) && (1)\\ &= \sum_{i=1}^{m}\varphi(\langle v, e_i\rangle e_i) && (2)\\ &= \sum_{i=1}^{m}\langle v, e_i\rangle \varphi(e_i) && (3)\\ &= \sum_{i=1}^{m}\langle v, \overline{\varphi(e_i)}e_i\rangle && (4)\\ &= \langle v, \sum_{i=1}^{m}\overline{\varphi(e_i)}e_i\rangle. && (5) \end{aligned}$$

To prove uniqueness, suppose $\varphi(v) = \langle v, u_1\rangle = \langle v, u_2\rangle$. Then by linearity in the second slot we have $0 = \langle v, u_1 - u_2\rangle$. Since this must hold for all $v \in V$, we can conclude $u_1 - u_2 = 0$ and $u_1 = u_2$. $\square$

Now that we've proved the Riesz Representation Theorem, we can explore many of its amazing implications in linear algebra. For the rest of these notes, I will refer to it as simply the Riesz Theorem.

1: The Riesz Theorem establishes an explicit (albeit constructed) isomorphism between V and V', where V' is the vector space dual to V. To see this, take V to be an inner product space, or a vector space with the intrinsic mapping $\langle \cdot, \cdot\rangle : V \times V \to F$. Now we can define a natural map

$v \mapsto \varphi_u$ where $\varphi_u(v) = \langle v, u \rangle$. This map sends every $v \in V$ to V'. φ_u is well defined because of the Riesz Theorem.

2: The Riesz Theorem plays an important role in showing that the orthogonal complement of a subspace U is the same as the annihilator on U translated back to V. Recall that the annihilator on U where $U \subset V$, denoted U^0, is defined as $U^0 = \{\varphi \in V' : \varphi(u) = 0 \; \forall u \in U\}$.

Proof.

$$
\begin{aligned}
U^\perp &= \{v \in V : \langle v, u \rangle = 0 \; \forall u \in U\} && (6)\\
&= \{v \in V : \langle u, v \rangle = 0 \; \forall u \in U\} && (7)\\
&= \{v \in V : \varphi_v = 0\} && (8)\\
&= \{v \in V : \varphi_v \in U^0\} && (9)\\
& && (10)
\end{aligned}
$$

□

The step from line 7 to line 8 is possible because of the Riesz Theorem. Interestingly, this result allows us to gain deeper insight as to why the left null space of a matrix A is orthogonal to the column space of A. The left null space of a matrix A is simply the null space of A^T. Consider T', the dual map to T where $T : V \to W$ and $T' : W' \to V'$. Recall that the matrix of T' with respect to some bases of V' and W' is A^T. Now, we will show that $nullT' = (rangeT)^0$.

Proof. Suppose $\varphi \in nullT'$. Thus $0 = T'(\varphi) = \varphi \circ T$, and $0 = (\varphi \circ T)(v) = \varphi(Tv)$ for every $v \in V$. This gives us $\varphi \in (rangeT)^0$, and hence $nullT' \subset (rangeT)^0$. To prove inclusion in the opposite direction, suppose $\varphi \in (rangeT)^0$. Thus $\varphi(Tv) = 0$ for every vector $v \in V$. Hence $0 = \varphi \circ T = T'(\varphi)$, or $\varphi \in nullT'$, which shows $(rangeT)^0 \subset nullT'$. □

We can now combine the two previous results: $nullT' = (rangeT)^0$, and because $U^\perp$ is essentially U^0 translated back to V, $(rangeT)^\perp$ is the same as $(rangeT)^0$ translated back to W, since $rangeT \subset W$. Thus, $nullT'$ is essentially $(rangeT)^\perp$, and it follows that the left null space of a matrix A is orthogonal to its column space.

Allow me to digress for a bit on this topic: the least squares solution to a system of linear equations is given by $x = (A^T A)^{-1} A^T y$. This solution often relies on $A^T A$ being an invertible matrix. In many engineering and machine learning applications of least squares, an *overdetermined* system, or one with more equations than constraints, is preferred to one that is not overdetermined precisely to make sure $A^T A$ is invertible. We can see that this makes sense because if the system is overdetermined, the matrix A will have more rows than columns, which means its columns will be linearly independent as long as the features chosen for the system are not all redundant. Now, the question is: why is $A^T A$ invertible if the columns of A are linearly independent? To see this, recall that this means the transformation of A is injective and consider the following:

$$
A^T A x = 0 \tag{11}
$$

Note that this equation tells us $Ax \in nullA^T$, or in other words, $colA \subset nullA^T$. Recall from above that $nullA^T \perp colA$. It isn't too hard to see that $nullA^T$ and $colA$ are orthogonal complements, and from our results above this means that $nullA^T$ annihilates $colA$ on W, where W

is the codomain of the transformation of A. Thus, we have that $Ax = 0$ and since $Ax = 0$ only if $x = 0$, $A^T Ax = 0$ only if $x = 0$.

Another (and perhaps more elegant) way to see this is to show that $nullA = nullA^T A$:

$$A^T Ax = 0 \iff x^T A^T Ax = 0 \iff (Ax)^T Ax = 0 \iff ||Ax||^2 = 0 \iff Ax = 0 \tag{12}$$