8/31/05

Linear Algebra Review

Independence, Spanning, and Dimension

Def/ A set of vectors $x^1, \ldots, x^k$ is said to be *linearly dependent* if there exists a vector $\lambda \neq 0$ such that $\sum_{i=1}^k \lambda_i x^i = 0$. Otherwise, the set is *linearly independent.*

Claim/ If S is linearly dependent and $S \subset T$, then T is also linearly dependent. If some set S is linearly independent and $S \supset T$, then T is also linearly independent.
Proof: If $T = S$, we're done. Otherwise, WLOG $T = x^1, \ldots, x^k$ and $S = x^1, \ldots, x^l$ where $l < k$. Since S is linearly dependent, $\exists \lambda \neq 0, \lambda \in \mathbb{R}^l$ such that $\sum_{i=1}^l \lambda_i x^i = 0$. However, letting $\lambda_{l+1} = \cdots = \lambda_k = 0$ gives $\sum_{i=1}^k \lambda_i x^i = 0$. Thus, T is linearly dependent. □
The second statement is just the contrapositive.

Def/ A (usually infinite) set of vectors S is a *vector space* if $\forall x, y \in S, \lambda \in \mathbb{R}$, (a) $x + y \in S$ (b) $\lambda x \in S$.
Def/ A set S is said to *span* a vector space V if all elements of V can be written as linear combinations of vectors in S. For a set of vectors T, *span*(T) is the set of all vectors that can be expressed as linear combinations of vectors in T, i.e. the largest vector space that T spans.

Claim/ For any set T, span(T) is a vector space.
Proof: Suppose x and y are linear combinations of vectors in T. Then $x + y$ and λx are also linear combinations of vectors in T. Thus, span(T) is a vector space. □

Def/ A set of linearly independent vectors S is a *basis* for a subspace V if $S \subset V$ and S spans V.
Ex/ The standard basis for $\mathbb{R}^n$ is the set $e^1, \ldots, e^n$ where e^i is the vector of zeros with 1 in the i^{th} position.

Proof: Assume that for some λ, $\sum_{i=1}^{n} \lambda_i e^i = 0$. Restricting attention to the j^{th} coordinate in this sum, we have that $\sum_{i=1}^{n} \lambda_i e^i_j = 0$, but e^i_j is only nonzero when $j = i$, which implies that $\lambda_i e^i_i = 0$. Thus, $\forall j, \lambda_j = 0$, which implies that $e^1, \ldots, e^n$ is linearly independent.

Now, consider some vector $x \in \mathbb{R}^n$. If we let $\lambda_j = x^j$, we have that $x = \sum_{i=1}^{n} \lambda_i e^i$. Thus, every vector in $\mathbb{R}^n$ is a linear combination of $e^1, \ldots, e^n$, so this set is a basis. $\square$

Claim/ If $S = \{x^1, \ldots, x^k\}$ is linearly dependent, then $\exists j$ such that x^j is a linear combination of $x^1, \ldots x^{j-1}$.
Proof: Since S is linearly dependent, $\exists \lambda \neq 0$ such that $\sum_{i=1}^{k} \lambda_i x^i = 0$. Let j be the largest index such that $\lambda_j \neq 0$. This implies that $\sum_{i=1}^{j} \lambda_i x^i = 0$. Since $\lambda_j \neq 0$, we can divide and obtain $x^j = \sum_{i=1}^{j-1} -\frac{\lambda_i}{\lambda_j} x^i$. Thus, x^j is a linear combination of $x_1, \ldots, x_{j-1}$. $\square$

Claim/ If S, T are linearly independent sets in vector space V, and S a basis for V, and $|S| = n, |T| = k$. Then $k \leq n$.
Proof: Assume $k > n$. Since S, T are linearly independent, no vector in S or T is zero. Let $S = \{x^1, \ldots, x^n\}, T = \{y^1, \ldots, y^k\}$. Since S spans V, y^1 is a linear combination of $\{x^1, \ldots, x^n\}$, which means the set $\{y^1, x^1, \ldots, x^n\}$ is linearly dependent. By the previous claim, there is some vector the is a linear combination of the previous vectors. It cannot be y^1, so it is some x^j. Let S_1 be $\{y^1, x^1, \ldots, x^n\}$ with x^j removed. Since S spans V and x^j is a linear combination of elements in S_1, we have that S_1 spans V as well.

Now, consider the set $\{y^2\} \cup S_1$. Again, this set is linearly dependent since S_2 spans V. If we order the elements of this set $\{y^1, y^2, x^1, \ldots, x^n\}$, we can apply the previous claim again, and we know that the resulting element must be some x^l since $\{y^1, y^2\}$ are linearly independent. So let $S_2 = S_1 \cup \{y^2\} - \{x^l\}$. Again, S_2 spans V.

We can continue this process, adding elements of T, always preserving the property that S_i spans V. However, since $k > n$, we will reach the set $S_n = \{y^1, \ldots y^n\}$ which spans V. However, y^{n+1} is in V, which means that it is a linear combination of $\{y^1, \ldots y^n\}$. This is a contradiction since T is linearly independent. Thus, $k \leq n$. $\square$

Cor/ All bases for a vector space V have the same cardinality.
Cor/ Any linearly independent set $S \subset \mathbb{R}^n$ has $|S| \leq n$.

Def/ The *dimension* of V is the size of any basis of V.

Matrices, Rank, and Invertibility

Def/ For a matrix $A \in \mathbb{R}^{m\times n}$, the *row (column) space* of A is the vector space spanned by its row (column) vectors. A has *row (column) rank* k if the basis of its row (column) space has size k.
Claim/ For any matrix A, the row rank and column rank of A are equal.
Proof: Consider some matrix $A \in \mathbb{R}^{m\times n}$. Assume its row rank is k, and that the set $\{y^1,\ldots,y^k\}$ is a basis for the row space. Then the i^{th} row $r^i = (a_{i1},\ldots,a_{in})$ can be expressed as $\sum_{r=1}^{k}\lambda_{ir}y^r$. Looking at the j^{th} co-ordinate of this sum, we have that $a_{ij} = \sum_{r=1}^{k}\lambda_{ir}y^r_j$. Since this is true for all j, we have that the j^{th} column $c^j = (a_{1j},\ldots,a_{nj})$ can be expressed as $\sum_{r=1}^{k} y^r_j z^r$, where $z^r = (\lambda_{1r},\ldots,\lambda_{nr})$. This means that every column of A is a linear combination of k vectors, which means that the column space can have dimension no larger than k. So column rank of $A \leq$ row rank of A.

Applying this argument to the column space gives the other inequality, that row rank of $A \leq$ column rank of A. Thus, row rank and column rank are the same. □

Def/ (Matrix multiplication) Given matrix $A \in \mathbb{R}^{m\times r}, B \in \mathbb{R}^{r\times n}$, we have that $C \in \mathbb{R}^{m\times n}$ is the product of A and B, denoted $AB = C$ if C is the matrix where $C_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj}$.

$$c_{21} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}\begin{pmatrix} b_{11} & \cdot \\ b_{21} & \cdot \\ b_{31} & \cdot \\ b_{41} & \cdot \end{pmatrix}$$

Matrix multiplication is associative and distributive, but **not** commutative.
Def/ A matrix $A \in \mathbb{R}^{n\times n}$ has inverse B if $AB = BA = I_n$, where

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix}$$

A matrix is *invertible* or *nonsingular* if it has an inverse. Otherwise it is *singular*. The inverse of a matrix A is denoted A^{-1}.
Claim/ If $A \in \mathbb{R}^{n\times n}$ has an inverse, it is unique.

Proof: Assume A has inverses B and C. Then consider $D = BAC$. Associating one way, this is

$$D = B(AC) = B(I_n) = B.$$

Associating the other way, this is

$$D = (BA)C = (I_n)C = C.$$

This implies that $B = C$, which implies that the inverse is unique. □

Fact/ If $AB = I_n$ then $BA = I_n$.

Claim/ If A, B invertible, then AB is invertible.
Proof: The inverse of AB is $B^{-1}A^{-1}$ since

$$B^{-1}A^{-1}AB = B^{-1}I_nB = B^{-1}B = I_n$$

Claim/ A matrix $A \in \mathbb{R}^{n\times n}$ is invertible iff it has rank n.
Proof: ($\Rightarrow$) Assume A has rank $k < n$, and inverse A^{-1}. Since it does not have rank n, the columns of A, $a^1, \ldots, a^n$ are dependent. Thus, there is a $\lambda \neq 0$ such that $\sum_{i=1}^n \lambda_i a^i = 0$, or in matrix notation, $A\lambda = 0$. Now consider $A^{-1}A\lambda$. We have $(A^{-1}A)\lambda = I_n\lambda = \lambda$ associating one way, but we also have $A^{-1}(A\lambda) = A^{-1}0 = 0$. This is a contradiction since $\lambda \neq 0$. Thus, it must be the case that A has rank n.

($\Leftarrow$) Assume A has rank n. Then the columns of A span $\mathbb{R}^n$. Thus, we can write any vector in $\mathbb{R}^n$ as a linear combination of the columns of A. Specifically, for any j, we can write e^j as some $\sum_{i=1}^n \lambda_i^j a^i$. Then if we let matrix B have columns $(\lambda^1, \ldots, \lambda^n)$, we see that $AB = I_n$. Thus, A is invertible. □

Solving Systems of Equations

Given matrix $A \in \mathbb{R}^{m\times n}$ and (column) vector $b \in \mathbb{R}^m$, it's often useful to be able to solve for a vector $x \in \mathbf{R}^n$ that satisfies $Ax = b$.

Def/ A matrix $A \in \mathbb{R}^{m\times n}$ with $m < n$ is said to have *full rank* if rank(A) $= m$.
Claim/ If A has full rank, then the system $Ax = b$ always has a solution.

Proof: Since A has full rank, it has column rank m, which means we can find m linearly independent columns of A. WLOG, let those columns be the 1^{st} m columns. Then the matrix B which consists of those m columns is invertible, and if we left multiply by B^{-1}, we see that the first m columns of $B^{-1}A$ are the identity matrix. Thus, if $y = B^{-1}b$, then one solution to $Ax = b$ is the vector $(y_1, \ldots, y_m, 0, \ldots 0)$, since this satisfies $B^{-1}Ax = B^{-1}b$.□

So we can sometimes guarantee that a solution to a system of equations exists. But how can we actually find such a solution?

Ex/ Find solutions to the system $Ax = b$ where

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix} b = \begin{pmatrix} 9 \\ 8 \\ 3 \end{pmatrix}$$

In order to determine the set of solutions x, we reduce the problem using elementary row operations.

Def/ The *elementary row operations* are:

(1) Scaling a row by some nonzero constant

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \mathbf{2} & \mathbf{4} & \mathbf{6} \\ 2 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix}$$

(2) Interchanging two rows

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \mathbf{2} & \mathbf{-1} & \mathbf{1} \\ \mathbf{1} & \mathbf{2} & \mathbf{3} \\ 3 & 0 & -1 \end{pmatrix}$$

(3) Adding some multiple of one row to another

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix} = \begin{pmatrix} \mathbf{5} & \mathbf{0} & \mathbf{5} \\ 2 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix}$$

Each elementary row operation corresponds to left multiplying by a certain square matrix, known as an *elementary matrix.*

Fact/ All elementary matrices are invertible.

Using this fact, we can apply elementary row operations to our matrix and RHS vector to simplify the problem. Since elementary matrices are invertible, we preserve the solution set.

Claim/ Let E be an invertible matrix. Then x satisfies $Ax = b$ iff x satisfies $EAx = Eb$.

Proof: If $Ax = b$, left multiplying by E gives $EAx = Eb$. If $EAx = Eb$, left multiplying by E^{-1} gives $Ax = b$. □

Now, we can define the *Gauss-Jordan Elimination* method for solving systems of equations. This method involves applying elementary row operations to the matrix and RHS vectors to reduce the problem to a simpler form.

Ex/ Solve $Ax = b$ for the previously defined A, b.

Start by augmenting the matrix with the RHS vector:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 2 & -1 & 1 & 8 \\ 3 & 0 & -1 & 3 \end{array}\right)$$

Since the first element of the first row is 1, eliminate all entries in the first column under that element.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & -5 & -5 & -10 \\ 0 & -6 & -10 & -24 \end{array}\right)$$

The second element of the second row is nonzero, but not 1, so scale that row.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 1 & 2 \\ 0 & -6 & -10 & -24 \end{array}\right)$$

Eliminate the rest of the second column

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & -4 & -12 \end{array}\right)$$

Scale the third row so that its leading nonzero is one.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 9 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Eliminate entries above the diagonal

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Thus, the unique solution to the original problem is $x = (2, -1, 3)$. Note that in this example the leading elements of the rows were not zero, so we could scale them to 1. If one of these elements was zero, we would either use the interchange operation to swap in a row that had a nonzero element in that position, or else continue to the next column since the current column had only zeros in active rows.