Department of Mathematics, University of Houston, Math 4389
Some intermediate linear algebra and basic matrix theory facts–Blecher

It is important first to establish the purpose of this document. It can be used as a review of some of the things you may not remember from linear algebra (although it does not include some early things that it is assumed that you know, like the meaning of the words in the first few bullets below). So drink it in, letting it settle and remind you (or inform you). Parts of it can be used as a 'formula sheet'. There is some repetition in parts of this document, usually there to help view some things from different angles. There are almost no examples (you do not find examples on a formula sheet), you will have to supply those from the usual sources when you need them. There may be some more simple things that are not said, since we are often focussing here on things that are harder to remember. In places there is a lot more than you need. However towards the end of this document the detail will tail off, because I ran out of energy and time, and be replaced by lists of topics a linear algebra student needs to know (fortunately most of these are just 'recipes'). I will probably add and change some things in this document later.

A nice thing about linear algebra is that from a higher math/proofs point of view it is a relatively easy subject–that is an able graduating math major at a good school should, theoretically, be able to do almost any of the proofs as exercises. Try that! (Not meant as an insult, but as a challenge, and to reassure you that the proofs are uasually easy. Of course its easier if you've taken Advanced Linear Algebra.)

- It is assumed you know your vectors in $\mathbb{R}^n$, dot product, their length (Euclidean norm $\|\vec{v}\|_2 = \sqrt{\vec{v} \cdot \vec{v}}$), angle between them, etc. Cauchy-Schwarz inequality: $|\vec{a} \cdot \vec{b}| \leq \|\vec{a}\|_2 \, \|\vec{b}\|_2$.
- You are expected to know the meaning of 'size of a matrix' (i.e. $m \times n$) and when two matrices have the same size, or when they are equal, or square ($m = n$).
- It is assumed you know your matrix algebra.
- $A\vec{x} = \sum_k x_k \, \vec{a}_k$, where $\vec{a}_k$ is the kth column of A, and x_k is the kth entry of $\vec{x}$. The matrix product
$$AB = [A\vec{b}_1 \; : \; A\vec{b}_2 \; : \; \cdots : A\vec{b}_r]$$
if $\vec{b}_k$ is the kth column of the $n \times r$ matrix B. Note the i-j entry of AB is the dot product of the ith row of A and the jth column of B.
- You are expected to know the rules for the transpose A^T, like $(AB)^T = B^T A^T$. We call A symmetric if $A = A^T$.
- You are expected to know a few basic things about *permutation matrices* (wiki). These are matrices of 0's and 1's, with exactly one 1 in every column and row.
- You are expected to know what is an upper or lower triangular or diagonal matrix (symbols U, L, D), and what the 'main diagonal' of a matrix is. You may be asked for the LU and maybe the LDU decomposition (wiki), although this is not likely.
- We write I or I_n for the $n \times n$ *identity matrix*. This is both a diagonal matrix and a permutation matrix.
- We will discuss the inverse A^{-1} of a square matrix A in more detail later, but for now: a matrix is *nonsingular* iff it is invertible (has an inverse, that is, a matrix B with

$AB = BA = I$. We write such B, if it exists, as A^{-1}). Otherwise it is *singular*. The inverse of a diagonal matrix $\text{diag}(d_1, d_2, \cdots, d_n)$ is $\text{diag}(d_1^{-1}, d_2^{-1}, \cdots, d_n^{-1})$, assuming the d_k are all nonzero. The inverse of a 2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

assuming $ad - bc \neq 0$. (If $ad - bc = 0$ the matrix is not invertible. For a general 3×3 or larger square matrix A you can find the inverse A^{-1} if it is invertible either by (1) the cofactor method using the determinant (wiki), or (2): Try to use Gauss-Jordan elimination (below) on the matrix $[A : I]$ to achieve reduced row echelon form $[I : C]$. If you can do this then $C = A^{-1}$, otherwise A is not invertible.

- If A is an $m \times n$ matrix, we can use *Gauss elimination* to *row reduce* A to *row echelon form (REF).* We sometimes say 'staircase' to refer to the final echelon form. We can also use *Gauss-Jordan elimination* to row reduce A to *reduced row echelon form* (RREF). By *an echelon form of* A we mean either of the above. We will write RREF(A) for the reduced row echelon form. Convention: we will make RREF(A) have 1's in all the pivots (so that each pivot column has one 1 and the rest zeros).
- One may view the row echelon form or the RREF as the product EA of a matrix E and A. Indeed E is the product of 'elementary matrices' each representing and performing (by left multiplying by it, in the right order) one of the *elementary row operations* needed in the row reduction. There are three elementary row operations, namely switching two rows, multiplying a row by a nonzero scalar k, or adding to one row a nonzero multiple of another row. Hence there are three kinds of elementary matrices, all invertible. The first is a simple permutation matrix (which is its own inverse), the second is a simple diagonal matrix which looks like the identity matrix except for a k in one of the diagonal entries (and its inverse is the same except except for a $1/k$ in that entry). The third kind, adding to one row a nonzero multiple k of another row is a lower triangular matrix which looks like the identity matrix except for a k in one of the spots below the main diagonal (and its inverse is the same except except for a $-k$ in that spot).
- You are expected to know the meaning of 'pivot positions', 'pivot columns' and 'free variables' (the variables corresponding to the non-pivot columns). Non-pivot columns are sometimes called 'free columns'.
- The *augmented matrix* of a linear system $A\vec{x} = \vec{b}$ is the matrix $[A : \vec{b}]$.
- We usually view vectors in $\mathbb{R}^n$ as columns (as in the last line). Sometimes to stress this point we call them column vectors.

- We will not define a vector space here. If you have not looked at that for a while do so now, eg. on wikipedia, and be sure that you have a feeling for which examples are vector spaces and which are not. The main vector spaces we look at in this course are $\mathbb{R}^n$. and its subspaces (particularly the ones met with below). By a *scalar* we mean a real number (although nearly everything here applies to complex vector spaces too).

- Recall that a *subspace* of a vector space V is a subset W of V containing the 0 vector of V, such that (a) $v+w \in W$, and (b) $cv \in V$, for all $v, w \in W$ and for all scalars c. FACT: Any subspace is a vector space.
- A *linear combination* of vectors $v_1, v_2, \cdots, v_m$ is a vector of form $c_1v_1+c_2v_2+\cdots+c_mv_m$ for scalars $c_1, \cdots, c_m$. Here $m \in \mathbb{N}$.
- Writing one vector $\vec{b}$ in $\mathbb{R}^n$ as a linear combination of several other vectors: Number these other vectors, and let A be the matrix having these vectors as columns in this order. Solve $A\vec{x} = \vec{b}$. The coefficients in the desired linear combination are the entries in any particular solution of this equation.
- Recall that the solution of $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$ if A is square and invertible.

- The *span* of a set B of vectors is the set of all linear combinations of vectors in B; it is written Span(B). FACT: Span(B) is a subspace. We say a set B of vectors in a vector space V *spans* V, or is a *spanning set* for V, if Span$(B) = V$.

- A set of vectors $B = \{v_1, v_2, \cdots, v_m\}$ in a vector space V is said to be *linearly dependent* if one of them can be written as a linear combination of the others. Otherwise, B is said to be *linearly independent* or *l.i.* for short.

  A set of one vector $\{v\}$ is always linearly independent. Two vectors are linearly dependent if and only if one is a scalar multiple of the other. Two vectors in $\mathbb{R}^n$ are linearly independent if and only if they are not parallel. Three vectors in $\mathbb{R}^3$ are linearly independent if and only if they do not lie in the same plane.
- FACT: $B = \{v_1, v_2, \cdots, v_m\}$ is linearly independent if and only if the only scalars $c_1, c_2, \cdots, c_m$ such that $c_1v_1 + c_2v_2 + \cdots + c_mv_m = 0$ are $c_1 = c_2 = \cdots = c_m = 0$.
- Test if a collection of several vectors in $\mathbb{R}^n$ is linearly independent: Let A be the matrix having these vectors as columns. If this (or its RREF) has free variables they are linearly dependent. Saying these vectors are linearly independent is the same as saying that $A\vec{x} = \vec{0}$ has only the trivial solution $\vec{0}$.

- A *basis* for a vector space V is a finite set in V which is both linearly independent and spans V.

- A finite set $B = \{v_1, v_2, \cdots, v_m\}$ in V is a basis for V iff every vector$v$ in V can be written as a linear combination $v = c_1v_1 + c_2v_2 + \cdots + c_mv_m$ in one and only one way. We call this one way writing (or expressing) v *in terms of the basis*.
- A vector space V is *finite dimensional* if it has a (finite) basis. In this case one can show that all bases of V are the same size. This size is called the dimension of V, and is written as dim(V). If a vector space V is not finite dimensional it is called *infinite dimensional* (and then we will not talk about bases for V).
- FACT: The dimension of the span of a set S of several vectors in $\mathbb{R}^n$ is the maximum number of linearly independent members of S.
- FACT: A strictly smaller subspace (that is $W \subset V$ but $W \neq V$) will have strictly smaller dimension.
- Suppose that $\dim(V) = n$. Any set of strictly more than (resp. less than) n vectors in V is linearly dependent (resp. cannot span V). Any linearly independent (resp.

spanning) set in V has $\leq n$ (resp. $\geq n$) elements. So any set of linearly independent vectors in $\mathbb{R}^n$ has $\leq n$ members, any spanning set for $\mathbb{R}^n$ has $\geq n$ members.

- FACT: A set of n vectors in $\mathbb{R}^n$ is linearly independent if and only if it spans $\mathbb{R}^n$. That is, if and only if it is a basis for $\mathbb{R}^n$. This is also equivalent to saying that the matrix with these vectors as columns is invertible (or that this matrix has nonzero determinant).
- The *standard basis* or *canonical basis* for $\mathbb{R}^3$ from Calculus III is often written as $\vec{i}, \vec{j}, \vec{k}$. The *standard basis* or *canonical basis* for $\mathbb{R}^n$ is often written as $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$.

- Column space $C(A)$ of a matrix A:
- If A is an $m \times n$ matrix, we can write $A = [\vec{a}_1 : \vec{a}_2 : \cdots : \vec{a}_n]$; here $\vec{a}_k$ is the kth column of A. Then $C(A)$ is the set of all the linear combinations of $\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n$. So we really have $C(A) = \text{Span}(\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n)$. This is a vector space, is a subspace of $\mathbb{R}^m$ as we said above.
- Also, $C(A) = \{\vec{b} \in \mathbb{R}^m : A\vec{x} = \vec{b}$ has a solution $\}$. That is, asking if $\vec{b} \in C(A)$ is the same as asking: does $A\vec{x} = \vec{b}$ have a solution?

  (To prove this, recall that for any $\vec{x} \in \mathbb{R}^n$, we can write $A\vec{x} = x_1\,\vec{a}_1 + x_2\,\vec{a}_2 + \cdots + x_n\,\vec{a}_n$ (here x_k is the kth entry of $\vec{x}$). Thus saying that $A\vec{x} = \vec{b}$ has a solution, is the same as saying that $\vec{b}$ is a linear combination of the columns of A, or equivalently, as saying that $\vec{b} \in \text{Span}(\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n)$.)
- So one way to find $C(A)$ is to use Gauss elimination to find the $\vec{b}$ such that $A\vec{x} = \vec{b}$ has a solution.
- Finding a 'nice' basis for $C(A)$: Do Gauss-Jordan to A^T to obtain RREF(A^T). Take the transpose of RREF(A^T). The nonzero columns of this, that is the nonzero columns of $(\text{RREF}(A^T))^T$, form a basis for $C(A)$.
- Finding a basis for $C(A)$ consisting of columns of A: Do Gauss-Jordan to A. The original columns of A standing in the pivot columns, are the desired basis. (Here is the idea for why this works, not that you care probably. If E is the matrix from page 2 that one must multiply A on the left by to get the echelon form, then because E is invertible, it follows that 'the columns that count' in A, are exactly 'the columns that count' in its echelon form. But 'the columns that count' for the column space of the echelon form matrix, are exactly the 'pivot columns'.)
- Finding a basis for the span of several vectors in $\mathbb{R}^n$: Number them. Let A be the matrix having these vectors as columns in this order, and apply one of the last two items to find a basis for $C(A)$.
- Given a finite set of vectors S find a subset which is a basis for Span(S): this is the same recipe as for the last item, but using the item before it to find a basis for $C(A)$.

- Row space $R(A)$ of a matrix A is the span of the rows of A. Note that $R(A) = C(A^T)$ and $C(A) = R(A^T)$.
- Finding a basis for $R(A)$: Do Gauss-Jordan to A. The nonzero rows in RREF(A) are a basis for $R(A)$. Or you could find a basis for $C(A^T)$ by one of the methods above.

- rank(A) = number of nonzero rows in RREF(A) = number of pivot columns/free variables = $\dim(C(A)) = \dim(R(A))$.

- FACT: $\text{rank}(A) = \text{rank}(A^T)$.
- A *rank 1 matrix* may also be described as a matrix of the form $A = \vec{v}\,\vec{w}^T$, for vectors $\vec{v}, \vec{w}$.

- The *nullspace* $N(A)$ for matrix A is the set of solutions to the homogeneous equation $A\vec{x} = \vec{0}$. It is a subspace, and always contains the zero vector $\vec{0}$, the trivial solution to $A\vec{x} = \vec{0}$. The nullspace is also the *kernel* Ker (L_A) of the operator $L_A : \mathbb{R}^n \to \mathbb{R}^m$ of left multiplication by A on (column) vectors in $\mathbb{R}^n$. Then $\text{nullity}(A) = \dim N(A)$.
- If A has more columns than rows then $A\vec{x} = \vec{0}$ has infinitely many solutions, and $\text{nullity}(A) \geq 1$.
- Finding a basis for $N(A)$: The *fundamental solutions* or *special solutions* are a 'nice' basis for $N(A)$. Here is the recipe for finding the fundamental solutions. First compute RREF(A). Then each fundamental solution is found by solving RREF$(A)\,\vec{x} = \vec{0}$, putting all the free variables equal to zero except one, which we set equal to 1. (It is 'nice' because of all the '1's in positions where all other vectors have 0's.)

  Thus the number of fundamental solutions is the number of free variables in an echelon form of A. This number is the dimension of $N(A)$. If $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ are the fundamental solutions (or indeed are any basis for $N(A)$) then the general solution to $A\vec{x} = \vec{0}$ is $c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_m\vec{x}_m$ for scalars $c_1, c_2, \cdots, c_m$.
- Rank-nullity theorem: $\text{rank}(A) + \text{nullity}(A) = n$ for an $m \times n$ matrix A.

  (Sketch proof: this is almost obvious for RREF(A), and none of the operations for RREF change rank, nullity, n.)

- Adding to a linearly independent set B in $\mathbb{R}^n$ to get a basis for $\mathbb{R}^n$: Let A be the matrix having these vectors as columns. Then add to B a basis for $N(A^T)$. (Hint at a proof: $(\text{Span}(B))^\perp = C(A)^\perp = N(A^T)$ (see later section on orthocomplements)).

- The general solution to $A\vec{x} = \vec{b}$ is

$$\vec{x} = \vec{x}_h + \vec{x}_p = \vec{x}_p + c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_m\vec{x}_m.$$

  Here $\vec{x}_h$ is the general solution to the associated homogeneous equation $A\vec{x} = \vec{0}$, and $\vec{x}_p$ is a particular solution to $A\vec{x} = \vec{b}$, and $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_m$ are the fundamental solutions to the associated homogeneous equation $A\vec{x} = \vec{0}$. One way to find a particular solution to $A\vec{x} = \vec{b}$ is to do Gauss-Jordan elimination to the augmented matrix, then set the free variables equal to 0.
- Similar fact in differential equations, which has the 'same' proof: Suppose that L is a linear differential operator, that x_h is the general solution to the associated homogeneous equation $Lx = 0$, and that x_p is a particular solution to $Lx = E$. Show that the general solution to $Lx = E$ is $x = x_p + x_h$. And $x_h = c_1x_1 + c_2x_2 + \cdots + c_mx_m$ for scalars $c_1, c_2, \cdots, c_m$, where $x_1, x_2, \cdots, x_m$ are the fundamental solutions, i.e. a basis for the solution space of the associated homogeneous equation.
- Table of the number of solutions to $A\vec{x} = \vec{b}$ for an $m \times n$ matrix A: If $m = n$ there is 1 solution if A is invertible; if A is not invertible then there are zero or infinitely many solutions. If $m > n = \text{rank}(A)$ there are 1 or no solutions. If $m < n$ there are none or infinitely many solutions (there are infinitely many if $\text{rank}(A) = m$). If

rank$(A) < \min\{n, m\}$ there are none or infinitely many solutions. (All of these can be seen/proved by thinking about RREF of the augmented matrix, and/or using the second last bullet).

- On inverses: If $AB = I$ or $BA = I$ and A and B have the same size then $A = B^{-1}$ (and $B = A^{-1}$).
- If A and B are invertible matrices of the same size then the inverse of AB is $B^{-1}A^{-1}$. If A is an invertible matrix with inverse B then the transpose A^T has inverse B^T. (To prove this you may use the fact that $(AC)^T = C^T A^T$ for matrices A, C.)
- Determinants: only make sense for square matrices.
- if A is a square matrix then A is invertible iff $\det(A) \neq 0$.
- You should know the Cramer's rule method for solving the system $A\vec{x} = \vec{b}$ if A is a square matrix.
- You are expected to know the main properties of determinants: like $\det(A) = \det(A^T), \det(A^{-1}) = 1/\det(A), \det(AB) = \det(A)\det(B)$, etc.
- Or, that the determinant changes sign if you switch two rows or columns, is multiplied by k if you multiply a row (or column) through by k, and is unchanged if you add to one row a nonzero multiple of another row . If one row (resp. column) of a matrix equals or is a multiple of another row (resp. column) its determinant is zero. Etc.
- You are expected to know the formula for the determinant of a 2×2, or of an upper or lower triangular matrix. You are expected to know the 4 methods for the determinant of a 3×3 matrix, three of which work for bigger square matrices too (cofactors, Gauss elimination to lower triangular, the permutation definition).
- You should know the connections between determinants and the cross product of vectors, or the volume of a parallelopiped.

Theorem If A is an $n \times n$ matrix the following are equivalent:

(1) A is invertible.
(2) $\det(A) \neq 0$.
(3) rank$(A) = n$.
(4) For every vector $\vec{b}$ the system $A\vec{x} = \vec{b}$ has a solution.
(5) For every vector $\vec{b}$ the system $A\vec{x} = \vec{b}$ has a unique solution.
(6) For some vector $\vec{b}$ the system $A\vec{x} = \vec{b}$ has a unique solution.
(7) The system $A\vec{x} = \vec{0}$ has only the trivial solution $\vec{x} = \vec{0}$.
(8) The columns of A are linearly independent.
(9) The columns of A span n-space.
(10) The rows of A are linearly independent.
(11) The rows of A span n-space.
(12) The reduced row echelon form of A is I_n.
(13) The (echelon forms) of A have n pivot positions/columns.
(14) There is a matrix B with $AB = I$.
(15) There is a matrix B with $BA = I$.

There is a similar list of equivalent conditions for an $m \times n$ matrix, for when rank $(A) = m$ etc (and another such list when rank $(A) = n$). If interested ask me for these.

At this point I ran out of energy, and so end with mostly a (partial) list of linear algebra topics the linear algebra student should know:

- Orthogonal vectors, Orthogonal subspaces, Orthogonal complements $S^\perp$ of a set S in $\mathbb{R}^n$, orthonormal basis, Projection P of a vector onto a line or subspace, Construct projection $P = \sum_k \vec{v}_k\, (\vec{v}_k)^T$ using an orthonormal basis $(\vec{v}_k)$ for the subspace, rotation matrix.
- $N(A)^\perp = C(A^T) = R(A)$, so $N(A) = C(A^T)^\perp = R(A)^\perp$. And $C(A)^\perp = N(A^T)$ and $C(A) = N(A^T)^\perp$.
- Finding the orthogonal complement of a set S of several vectors in $\mathbb{R}^n$: Let A be the matrix having these vectors as columns. $S^\perp = N(A^T)$, since $(\mathrm{Span}(S))^\perp = C(A)^\perp = N(A^T)$.
- Application: Supplementing a linearly independent set B in $\mathbb{R}^n$ to a basis in $\mathbb{R}^n$: Let A be the matrix having these vectors as columns. Then add to B a basis for $N(A^T)$. Hint at a proof: $(\mathrm{Span}(B))^\perp = C(A)^\perp = N(A^T)$.
- Definitions of: Eigenvalue, Eigenvector, Characteristic Polynomial. The eigenspace corresponding to an eigenvalue c is the set of all vectors v with $Tv = cv$; that is, the nullspace of $A - cI$. Thee dimension of the eigenspace for an eigenvalue c is called the *geometric multiplicity* of that eigenvalue. The *algebraic multiplicity* of c is its order as a root of the characteristic polynomial (that is, the number of times that root is repeated in the polynomial). Trace of a Matrix. A symmetric real matrix A is *positive definite* if $(A\vec{x}) \cdot \vec{x} > 0$ for all nonzero vectors $\vec{x}$. It is *positive semidefinite* if $(A\vec{x}) \cdot \vec{x} \geq 0$ for all nonzero vectors $\vec{x}$. A symmetric real matrix is positive definite (resp. positive semidefinite) if all its leading principal minors (i.e. minor leading determinants) are > 0 (resp. ≥ 0). It is also the same as saying that all of its eigenvalues are > 0 (resp. ≥ 0). Relation of positive definite matrices to ellipses and hyperbolae and their major axes. Two square matrices A and B are *similar* if $A = S^{-1}BS$ for some invertible matrix S.

  Definitions of: Linear transformation, Kernel, Range, Matrix of a linear transformation, Composition of linear transformations, Inverse of a linear transformation, Identity transformation, Change of basis matrix.
- Let A be a matrix with real entries. If $\overrightarrow{v}$ is an eigenvector corresponding to a complex eigenvalue λ of A, let $\overrightarrow{v}^*$ be the vector obtained from $\overrightarrow{v}$ by taking the complex conjugate of each entry. Then $\overrightarrow{v}^*$ is an eigenvector corresponding to the eigenvalue $\bar{\lambda}$.
- Tasks/Exercises you must be able to do, related to the last definitions: Find the characteristic polynomial of a matrix, Find eigenvalues of a matrix, and associated eigenvectors, eigenspace, Eigenvalues of the inverse, or of powers, of a matrix, Cayley-Hamilton theorem, Checking if a matrix is positive definite, Least squares method, Using facts about similar matrices. Check if a function T is linear, Find kernel and range, Finding the matrix of a linear transformation. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear then $T = L_A$, left multiplication by A, where A is the $m \times n$ matrix whose jth column is $T(\vec{e}_j)$. Indeed this matrix A is $M_B(T)$ where B is the canonical basis $(\vec{e}_j)$. Computing

compositions and inverses of linear transformations, and their matrices, Computing a change of basis matrix.

- Some random remarks on diagonalization: We say that a square matrix A is *diagonalizable* (or can be diagonalized) if there is a matrix K of the same size such that $K^{-1}AK$ is diagonal. An $n \times n$ matrix can be diagonalized iff it has n linearly independent eigenvectors. Suppose that A is an $n \times n$ matrix, with n linearly independent eigenvectors $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$, and that λ_k is the eigenvalue associated with the eigenvector $\vec{v}_k$. Let K be the matrix $[\vec{v}_1 : \vec{v}_2 : \cdots : \vec{v}_n]$. Then

$$K^{-1}AK = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

  The last matrix is called a *diagonal* matrix.
- Suppose that $\lambda_1, \lambda_2, \cdots, \lambda_n$ are all the eigenvalues of an $n \times n$ matrix A, and suppose that they are all *distinct* (that is, there are no repeated eigenvalues). If $\vec{v}_k$ is *any* eigenvector corresponding to λ_k, for $k = 1, 2, \cdots, n$, then the set $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ is linearly independent.
- For any nonzero vector $\vec{v}$ we can obtain a vector of length 1, namely $\frac{1}{\|\vec{v}\|}\vec{v}$. This is called *normalizing* the vector.
- An *orthonormal set* of vectors is a set of vectors $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ such that each of these vectors has length 1, and $\vec{v}_i \cdot \vec{v}_j = 0$ for all i, j with $i \neq j$.
- An orthonormal set of vectors is linearly independent. An orthonormal set which is also spanning is an *orthonormal basis.*
- Recall that a matrix A is called *symmetric* if $A = A^T$. It is called *orthogonal* if $A^{-1} = A^T$.
- In $\mathbb{R}^n$, a set of vectors $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ is orthonormal if and only of the matrix $K = [\vec{v}_1 : \vec{v}_2 : \cdots : \vec{v}_n]$ is orthogonal.
- If A is a symmetric $n \times n$ matrix, then you can find an orthonormal basis $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ consisting of eigenvectors of A. Thus, by the facts above, if A is a symmetric $n \times n$ matrix, then there is an orthogonal matrix U such that $U^TAU = U^{-1}AU$ is diagonal. This is called *orthogonal diagonalization.* Indeed the columns of U are the members of the orthonormal set of eigenvectors of A above, and the diagonal matrix obtained has the corresponding eigenvalues on the diagonal.
- In some cases to find an orthonormal set of eigenvectors of A, one might need to use the *Gram-Schmidt process.* The Gram-Schmidt process constructs an orthonormal set from a set S of vectors (and it will be an orthonormal basis) if S is spanning.
- Another way to view diagonalization: picking a new basis of n-space such that (the matrix of) A with respect to this basis, becomes a diagonal matrix.