ME 8213 - Engineering Analysis - I
Fall 2017
Lecture Summary

Part II

Sep. 26 – Tuesday
LECTURE 12: Vector Spaces, Linear Combination, Linear Independence, Span, Basis, and Cartesian Products.

Summary:

1. The midterm exam will take place on Tuesday, Oct. 3. The midterm exam will be closed book, closed notes, no calculator, no computer, no cellphone, no smart watch. Only pencils and erasers are allowed. The exam will be based on the Exam Questions posted at the end of each lecture summary up to and including Lecture 11.
2. Review of the proof from Lecture 11 demonstrating that every linear function between two finite dimensional spaces can be expressed as a matrix multiplication.
3. A sheet of paper with x and y coordinates drawn on it is the main idea behind the concept of "vector space". The points on the sheet are vectors that fill up the "space" of the paper. The "space" so described is nothing else than a plane with the origin of the x-y coordinates included on it.
4. If one thinks of a vector space as a plane with a coordinate system, then it is seen that the applications of a plane are the same as the ones of a vector space. Planes are used to describe constraints or surfaces and are used to find component of forces in directions of interest, etc.
5. Higher dimensional spaces will be analogous to inserting additional independent coordinates to increase the amount of "space".
6. A two-dimensional plane drawn in three dimensions and passing through the origin is called a "subspace" because it is a smaller space (2D plane) inside a bigger space (3D). "Subspaces" are spaces that lie within bigger spaces. For example, two independent vectors can be used to form a coordinate system that defines a plane; however, if each of the vectors has five elements then the plane will be embedded in five dimensions, i.e., the plane containing the origin would be a 2D subspace embedded in 5D. This idea was discussed before in Lecture 2, when a circle was hidden (embedded) in five dimensions.
7. In order to properly define the concept of space we need to find a way to clearly and succinctly describe all points in the space. It is very difficult to do this in a few words so a good alternative is to define a few points and to explain how the rest of the points are generated. First consider two well-known operations that will become intrinsic to the definition of vector space: vector addition $(+)$ and scalar multiplication $(\cdot)$. In order to multiply a vector times a scalar we need to include the set of scalars in the definition of the vector space. This is commonly called a "field of scalars". The main concept behind the word "field" is that scalars can be added and multiplied together to form other scalars as well. A succinct description of a vector space is then:

"A vector space is a set of objects called vectors, along with a field of scalars and two operations, vector addition and scalar multiplication, such that space is closed both under vector addition and under scalar multiplication."

The phrase "closed under vector addition" indicates that adding two vectors in the space will result in a third vector in the space. Similarly "closed under scalar multiplication" indicates that multiplying a vector times a scalar results in another vector in the same space. In other words, it is not possible to escape the "space" using vector addition or scalar multiplication.

8. A *linear combination* of vectors $v^{\langle 0\rangle}, v^{\langle 1\rangle}, \cdots, v^{\langle N-1\rangle}$ is the "weighted" sum $c_0 v^{\langle 0\rangle} + c_1 v^{\langle 1\rangle} + \cdots + c_{N-1} v^{\langle N-1\rangle}$ where the weights are the scalars $c_0, c_1, \cdots, c_{N-1}$. Note that for column vectors, the linear combination is just a multiplication of the form

   $[v^{\langle 0\rangle} \quad v^{\langle 1\rangle} \quad \dots \quad v^{\langle N-1\rangle}] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix} = Ax$ where $A = [v^{\langle 0\rangle} \quad v^{\langle 1\rangle} \quad \dots \quad v^{\langle N-1\rangle}]$ is a matrix and $x = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-1} \end{bmatrix}$ is a vector containing the weights of the linear combination.

9. The equation $Ax = b$ is often interpreted as b being a linear combination of the columns of A, where the weights of the linear combination are the elements of vector x.
10. A set of vectors $\{v^{\langle 0\rangle} \quad v^{\langle 1\rangle} \quad \dots \quad v^{\langle N-1\rangle}\}$ is *linearly independent* if it is not possible to express one of the vectors in terms of the remaining vectors in the set. This implies that

    $c_0 v^{\langle 0\rangle} + c_1 v^{\langle 1\rangle} + \cdots + c_{N-1} v^{\langle N-1\rangle} \neq 0$

    for any arbitrary choice of weights $\{c_0, c_1, \cdots, c_{N-1}\}$ unless all weights are all equal to zero, i.e., there is no linear combination of vectors in the set that is identically equal to zero.
11. Using matrix notation, if $Ax \neq 0$ unless $x = \bar{0}$. Then the columns of matrix A must be linearly independent.
12. When a space is obtained by allowing all possible linear combination of an initial set of vectors, it is said that this initial set of vectors ***span*** the space.
13. A ***basis*** is a linearly independent set of vectors used to span a space.
14. It takes just a few sentences (try it!) to show that all possible *bases* (plural of basis) for a given vector space have the same number of vectors. This number is called the *dimension* of a vector space.
15. The **Cartesian Product** $X \times Y$ of subspaces X and Y is another vector space consisting of the set of all orders pairs $\{(x, y)\}$ so that, x is a vector of the space X and y is a vector of the space Y, i.e., $x \in X$ and $y \in Y$. For instance, if X and Y are each one dimensional spaces, then X and Y can be thought as the x and y axes of a plot. In this case, the X and Y consist of very thin lines that cannot possibly cover a two dimensional plane. However, the Cartesian Product of X and Y completely fills a two dimensional plane. The main idea here is that two very thin subspaces can be used to create a much larger subspace. The point (x, y) in vector space $X \times Y$ is often expressed as $(x, y) = x\hat{\imath} + y\hat{\jmath}$ where $\hat{\imath}$ and $\hat{\jmath}$ are unit vectors in X and Y, respectively.
16. Similarly, it is possible to create a Cartesian product of a two dimensional space and a one dimensional space. For example $(X \times Y) \times Z$ are the set of all orders pairs $\big((x, y), z\big) = (x, y, z)$, i.e., the coordinate system for three dimensions.

Exam Questions:

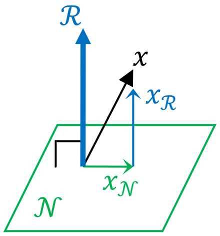
1. Provide a thorough explanation on how a vector space is defined.
2. Provide a thorough explanation on what is a linear combination of vectors.
3. Provide a thorough explanation on how to find if set of vectors of the same size (with the same number of elements) is linearly independent?
4. Provide a thorough explanation on what is meant when it is said that a set of vectors <u>span</u> a vector space?
5. Provide a thorough explanation on what is a basis for a vector space?
6. Is any plane a vector space? Explain your answer.
7. Let f1(x) = 1+x and f2(x) = 1-x be two independent, infinite-dimensional vectors defined in the interval [−1,1]. Find the component of the function g(x)=x^2 perpendicular to the subspace <u>spanned</u> by f1(x) and f2(x). Hint: any two vectors spanning the same space can be used as the basis for the space, can you find two independent linear combinations of f1 and f2 that span the same subspace and are much simpler functions? This will greatly simplify your answer.
8. Let the set of column vectors {x_1, x_2} represent a basis for R^2 and let {y_1, y_2 } be another set of column vectors representing a different basis for R^2. Given (a_1,a_2) such that z = a_1x_1 + a_2x_2, find (b_1,b_2) such that z = b_1y_1 + b_2y_2 for the same z. Note: the vectors forming the basis are not necessarily mutually orthogonal. You may use matrix equations.
9. What is the Cartesian product of two spaces?

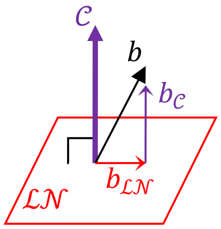
Sep. 28 – Thursday
LECTURE 13: The Four Spaces of a matrix.

Summary:

1. A matrix have four subspaces: two in the domain and two in the co-domain. In the domain we have two orthogonal subspaces: the **<u>nullspace</u>** and the **<u>rowspace</u>**. The nullspace $\mathcal{N}$ of a nxm matrix A (n rows, m columns) consists of the set of all $mx1$ vectors x such that $Ax = 0$. For this to be true, each vector x in $\mathcal{N}$ must be perpendicular to <u>all</u> the rows of A.
2. Proof that that the nullspace $\mathcal{N}$ of a matrix A is indeed a vector space: Let x_1 and x_2 be any two vectors in $\mathcal{N}$. Recall that a vector x is in $\mathcal{N}$ if $\mathrm{Ax} = 0$. This means that $\mathrm{Ax}_1 = 0$ and $\mathrm{Ax}_2 = 0$. It follows that $\mathcal{N}$ is a vector space if for any two arbitrary scalars α and β, $\mathrm{x}_3 = \alpha\mathrm{x}_1 + \beta\mathrm{x}_2$ is also part of $\mathcal{N}$. This is equivalent as showing that $\mathrm{Ax}_3 = 0$:
   $\mathrm{Ax}_3 = \mathrm{A}(\alpha\mathrm{x}_1 + \beta\mathrm{x}_2) = \alpha\mathrm{Ax}_1 + \beta\mathrm{Ax}_2 = \alpha 0 + \beta 0 = 0.$
3. The rowspace $\mathcal{R}$ of a matrix A is a vector space by definition because the rowspace is defined to be the set of all linear combinations of the rows of A and all linear combinations of a set of vectors comprises a vector space.
4. The columnspace $\mathcal{C}$ of a matrix A is defined to be the set of all linear combinations of the columns of A. Again, by definition all linear combinations of a set of vectors comprise a vector space, so $\mathcal{C}$ is a vector space.
5. It is important to realize that the dimension of the domain is the same as the number of elements in each vector in the domain. For instance, if the domain is $\mathbb{R}^3$, then each vector in it will have three elements. The same logic applies to the co-domain.
6. The columnspace $\mathcal{C}$ of a matrix A is the <u>range</u> of the linear function $f(x) = Ax$. In this case, the range is a subspace of the co-domain: the range is the part of the co-domain that is accessible through the linear function $f(x) = Ax$.

7. The left-nullspace $\mathcal{LN}$ of a matrix A is defined to be the set of all vectors perpendicular to the columns of A (and to all vectors in $\mathcal{C}$). Observe that if x is perpendicular to all the columns of A then $x^T A = 0$. Since x^T is to the <u>left</u> of A, and since $x^T A = 0$, i.e., the right hand side is zero or "<u>null</u>", this space is called the Left-Null space. It can be shown that $\mathcal{LN}$ is indeed a vector space by first selecting two vectors, say x_1 and x_2, in $\mathcal{LN}$. Since ${x_1}^T A = 0$ and ${x_2}^T A = 0$, it follows that if $\mathcal{LN}$ is indeed a vector space, then for any two arbitrary scalars α and β, $x_3 = \alpha x_1 + \beta x_2$ must also be part of $\mathcal{LN}$, i.e., ${x_3}^T A = 0$. This is easily demonstrated as follows: ${x_3}^T A = (\alpha x_1 + \beta x_2)^T A = \alpha {x_1}^T A + \beta {x_2}^T A = \alpha 0 + \beta 0 = 0$.
8. The domain of the function or mapping $f(x) = Ax$ is the cartesian product of $\mathcal{N}$ and $\mathcal{R}$, i.e., $\mathcal{N} \times \mathcal{R}$. Any vector x in the <u>domain</u> of $f(x) = Ax$ can be expressed in terms of a component in $\mathcal{N}$ and another component in $\mathcal{R}$, i.e., $x = (x_{\mathcal{N}}, x_{\mathcal{R}})$ or, since $x_{\mathcal{N}}$ is perpendicular (independent) to $x_{\mathcal{R}}$, the cartesian product can be expressed as $x = x_{\mathcal{N}} + x_{\mathcal{R}}$.





9. The co-domain of the function or mapping $f(x) = Ax$ is the cartesian product of $\mathcal{LN}$ and $\mathcal{C}$, i.e., $\mathcal{LN} \times \mathcal{C}$. Any vector b in the co-domain of $f(x) = Ax$ can be expressed as $b = b_{\mathcal{LN}} + b_{\mathcal{C}}$.
10. Since $f(x_{\mathcal{N}}) = Ax_{\mathcal{N}} = 0$, it follows that the entire nullspace of A is mapped into the origin of the co-domain.
11. The number of elements of each column is not necessarily the dimension of the columnspace, however it is equal to the dimension of the co-domain. The columnspace is a subspace of the co-domain.
12. The number of columns in a matrix is the dimension of the domain (which contains the rowspace). The number of rows of a matrix is the dimension of the codomain (which contains the columnspace).

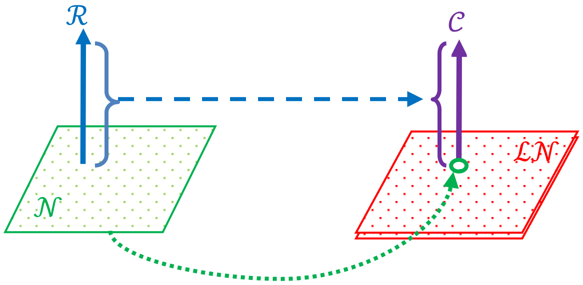
Exam Questions:

1) What is the Cartesian product of two spaces.
2) Given a nxm matrix A (n rows, m columns), prove that the set of all $mx1$ vectors x such that $Ax = 0$ make up a vector space.
3) Provide comprehensive definitions for the following spaces of a matrix: rowspace, nullspace, column space, and leftnullspace.
4) For a given nxm matrix A, explain why a vector in the rowspace would be perpendicular to a vector in the nullspace.
5) Prove that that the nullspace $\mathcal{N}$of a matrix A is indeed a vector space.
6) Prove that that the left-nullspace $LN$of a matrix A is indeed a vector space.
7) Prove that any vector in the domain of a nxm matrix A can be expressed as the sum of a vector in the rowspace and another vector in the nullspace.
8) Provide an explanation on why a matrix with a non-trivial nullspace (a non-trivial space is one that contains more than the zero vector), cannot be inverted.

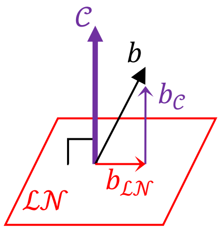
Oct. 10 – Tuesday
LECTURE 14: The Four Spaces of a matrix – continued

Summary:

1. The graded midterm exam was returned.
2. Review of previous lecture. Review of definitions of rowspace, columnspace, nullspace, and left-nullspace.
3. In the problem $Ax = b$ one can write $A(x_{\mathcal{N}} + x_{\mathcal{R}}) = b$ and since $Ax_{\mathcal{N}} = 0$, it follows that $Ax_{\mathcal{R}} = b$. It can be shown that for every b there is a unique vector $x_{\mathcal{R}}$ in $\mathcal{R}$ such that $Ax_{\mathcal{R}} = b$ as follows. Suppose that for a given b someone claimed to have found vectors $x_{\mathcal{R}1}$ and $x_{\mathcal{R}2}$ in $\mathcal{R}$ such that $Ax_{\mathcal{R}1} = b$ and $Ax_{\mathcal{R}2} = b$. Subtracting the latter equation from the former results in $A(x_{\mathcal{R}1} - x_{\mathcal{R}2}) = 0$. This means that $(x_{\mathcal{R}1} - x_{\mathcal{R}2})$ is in $\mathcal{N}$. This is not possible since, by definition, every linear combination of $x_{\mathcal{R}1}$ and $x_{\mathcal{R}2}$ must be in the rowspace of A unless it is the zero vector, i.e., the origin of the coordinate system. It follows that $(x_{\mathcal{R}1} - x_{\mathcal{R}2}) = 0$ or $x_{\mathcal{R}1} = x_{\mathcal{R}2}$, implying that for each b there can be only one vector $x_{\mathcal{R}}$ such that $Ax_{\mathcal{R}} = b$. Note that the converse is also true: For each $x_{\mathcal{R}}$ in $\mathcal{R}$ there is a unique b in $\mathcal{C}$ that cannot be reached through any other $x_{\mathcal{R}}$ in $\mathcal{R}$.



4. Since the domain of the function $f(x) = Ax$ is given by the Cartesian product $\mathcal{N} \times \mathcal{R}$**,** and the nullspace $\mathcal{N}$ is mapped into the origin (0) of the <u>co-domain</u> while all of the rowspace $\mathcal{R}$ is mapped, in a one-to-one manner, onto the columnspace $\mathcal{C}$, there is no way to reach the left-nullspace $\mathcal{LN}$ through the equation Ax. In other words, if $b_{\mathcal{LN}}$ is a vector in $\mathcal{LN}$ then $Ax \neq b_{\mathcal{LN}}$.
5. If b is in the co-domain of $f(x) = Ax$ then $b = b_{\mathcal{LN}} + b_{\mathcal{C}}$. Now, if $b_{\mathcal{LN}} \neq 0$, Ax cannot be equal to b. However, a very popular compromise is to solve the problem $Ax = b_{\mathcal{C}}$ instead as follows.



Since $b_{\mathcal{LN}}$ and $b_{\mathcal{C}}$ are orthogonal to each other, Pythagoras theorem applies. Pythagoras theorem is a sum of the square of the length of two orthogonal components. The "least squares" method consists in eliminating the squared term corresponding to $b_{\mathcal{LN}}$. The "least" part of "least squares" comes from the fact that the smallest change that can be made to b so that it lies on $\mathcal{C}$ is to truncate the component orthogonal to $\mathcal{C}$, i.e., eliminate $b_{\mathcal{LN}}$. This could be done by applying the Gram-Schmidt methodology to the columns of A and then using the resulting orthogonal unit vectors to find $b_{\mathcal{C}}$, i.e., the part of b parallel to $\mathcal{C}$.

6. Four problems related to the matrix equation $Ax \cong b$:
   a) Invertible (perfectly constrained or statically determinate): $Ax = b$
   b) Over constrained (Inconsistent or over-determined): $Ax \neq b$
   c) Under constrained (nullspace - under-determined, with infinite solutions): $Ax = b$
   d) Partly Under constrained and partly inconsistent: $Ax \neq b$

   a) is solved by inverting matrix A to obtain: $x = A^{-1}b$.
   b) can be addressed through the "least-squares" compromise: $Ax = b_{\mathcal{C}}$
   c) can be addressed through the "shortest solution" optimization: $Ax_{\mathcal{R}} = b$.
   d) can be addressed using both ***"shortest solution" and "least squares"***: $\mathrm{A}x_{\mathcal{R}} = b_{\mathcal{C}}$.

Exam Questions:

1. The domain of the function $f(x) = Ax$ is the Cartesian product of what two spaces? What can be said of the components of vector x along each of these spaces?
2. The co-domain of the function $f(x) = Ax$ is the Cartesian product of what two spaces? Consider the problem $Ax \cong b$, what can be said of the components of vector b with respect to these two spaces?
3. True or False: The nullspace and left-nullspace of a matrix have the same dimension. Clearly explain your answer.
4. Write down a matrix bigger than 1x1 with no nullspace and no left-nullspace. Explain your answer.
5. Write down a matrix with a two dimensional codomain, a two dimensional domain, and a one- dimensional left-nullspace.
6. Write down a matrix with no nullspace (except from the zero vector), a three dimensional codomain, and a two dimensional left-nullspace. Be able to answer different variations of this kind of problem.
7. In the problem $Ax = b$, prove that for every b there is a unique vector $x_{\mathcal{R}}$ such that $Ax_{\mathcal{R}} = b$. Clearly explain your answer and do not assume that A is invertible.

Oct. 12 – Thursday
LECTURE 15: The Four Spaces of a matrix – continued

Summary:

1. If there are r independent vectors $\{b_1, b_2, \cdots, b_r\}$ spanning the column space of A, the corresponding vectors $\{x_{\mathcal{R}_1}, x_{\mathcal{R}_2}, \cdots, x_{\mathcal{R}_r}\}$ in the rowspace will also be independent and will span

the row space: **the columnspace and rowspace have the same number of independent vectors.**

Proof: The proof can be split in two parts: 1) given r independent vectors spanning the columnspace of A, the corresponding r vectors in the rowspace will be shown to be independent, and 2) these r independent vectors in the rowspace do span the rowspace.

Part 1) Let $Ax_{\mathcal{R}_1} = b_1, Ax_{\mathcal{R}_2} = b_2, \ldots, Ax_{\mathcal{R}_r} = b_r$. These equations can be joined together into a single matrix equation as follows:

$$\underbrace{[b_1 \quad b_2 \quad \cdots \quad b_r]}_{B} = A\,\underbrace{[x_{\mathcal{R}_1} \quad x_{\mathcal{R}_2} \quad \cdots \quad x_{\mathcal{R}_r}]}_{M}$$

Observe that the columns of matrix B correspond to $\{b_1, b_2, \cdots, b_r\}$ and the columns of matrix M be the corresponding vectors $\{x_{\mathcal{R}_1}, x_{\mathcal{R}_2}, \cdots, x_{\mathcal{R}_r}\}$ in the rowspace of A. It follows that $B = AM$. Now consider $Bz = AMz$, where z is a nonzero and otherwise arbitrary vector of appropriate size. Since the left side of the equation, i.e., Bz, cannot be zero (because the columns of B are linearly independent) it follows, by the equality sign, that the right-hand side of the equation, i.e., AMz, cannot be zero, implying that Mz cannot be zero. This, combined with the fact that z is arbitrary implies that the columns of M must be linearly independent.

Part 2) Next we need to show that the columns of M span $\mathcal{R}$, i.e., that any arbitrary vector x' in $\mathcal{R}$ can be expressed as a linear combination of the columns of M. Consider an arbitrary vector x' in $\mathcal{R}$ with the corresponding vector b' in $\mathcal{C}$, i.e.,

(a) $b' = Ax'$

Recalling that any vector b' in the column space of matrix A can be expressed as $b' = Bz'$ for some z', and since, from part 1, $Bz' = AMz'$, so it follows that

(b) $b' = AMz'$

Subtracting (b) from (a) yields

(c) $0 = Ax' - AMz' = A(x' - Mz')$

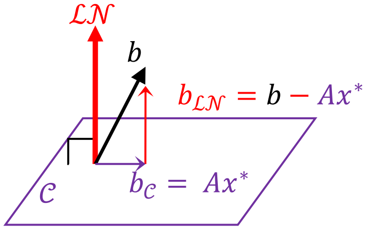
Since, both x' and Mz' are in the rowspace of A, the only way that $A(x' - Mz') = 0$ is that $(x' - Mz') = 0$ or,

(d) $x' = Mz'$

implying that the columns of M span $\mathcal{R}$ since any random vector x' selected from $\mathcal{R}$ is a linear combination of the columns of M.

2. The number of independent vectors in the columnspace (same number as in the rowspace) is called the **rank** of a matrix.
3. Below, it is shown how to project a vector in the co-domain onto the columnspace of a matrix A (assuming A to have independent columns.) First consider the right angle triangle obtained from the equation $b = b_{\mathcal{LN}} + b_{\mathcal{C}}$. Since $b_{\mathcal{C}}$ is in the columnspace of A one can write $Ax^* = b_{\mathcal{C}}$. Here Ax^* represents a unique linear combination of the columns of A - the weights or coefficients of the linear combination are stacked as the elements of x^*. It follows that $b_{\mathcal{LN}} =$

$b - b_{\mathcal{C}} = b - Ax^*$. Since $b_{\mathcal{LN}}$ is perpendicular to every vector in the column space of A, it must be perpendicular to all the columns of A, i.e., $A^T b_{\mathcal{LN}} = 0$ or substituting for $b_{\mathcal{LN}}$, $A^T(b - Ax^*) = A^T b - A^T Ax^* = 0$.



4. The equation $A^T b - A^T Ax^* = 0$ can be written as $A^T Ax^* = A^T b$. This latter equation is known at the "normal" equation.
5. If the columns of A are independent of each other, it can be shown that $A^T A$ is invertible as follows. First consider a non-trivial linear combination of the columns of A, i.e., Ax where $x \neq 0$. Since A has independent columns, then Ax is a nonzero vector in the column space A, implying that it cannot be orthogonal to all the columns of A, i.e.,
   $A^T(Ax) = A^T Ax \neq 0$.
   This means that $A^T Ax$ cannot be zero for an arbitrary, nontrivial x. It follows that $A^T A$ is a square matrix with independent columns, i.e., a full rank invertible matrix.
6. A set of independent columns spanning the columnspace of a matrix B can be obtained by using the row reduction echelon form (rref command in Mathcad) on B^T. You can read about row reduction echelon form in your textbook. Basically, row reduction performs the same steps as Gaussian Elimination. By combining rows to produce new rows, row reduction does not change the rowspace of matrix B. The non-zero rows of the result of $rref(B)$ constitute a basis for the rowspace of B. The resulting independent rows can be transposed and used as the columns of a matrix A than can be used to construct a projection matrix into the columnspace.
7. A better way to find an independent set of columns spanning the columnspace of B is to use the Gram-Schmidt Orthogonalization procedure on the columns of B and use the results to create a matrix A. In this case $A^T A = I$ because the columns of A will be orthogonal to each other and of unity length. One can use the qr decomposition in Mathcad, along with the submatrix command to obtain a Gram-Schmidt orthogonalization.
8. Assuming that the columns of A are independent of each other, it is possible to solve the normal equation for x^* to obtain $x^* = (A^T A)^{-1} A^T b$.
9. Using x^*, one can write $b_{\mathcal{C}} = Ax^* = A(A^T A)^{-1} A^T b$. This means that to project a vector b onto the column space of A, it is only necessary to multiply b by the "projection matrix"
   $P_{\mathcal{C}} = [A(A^T A)^{-1} A^T]$.
10. A good way to remember the formula for $P_{\mathcal{C}}$ is as follows. First consider the problem $Ax \neq b$ where A is a nxm matrix with n independent columns, then multiply by A^T to obtain $A^T Ax^* = A^T b$. Observe that the inequality went away because $A^T b = A^T(b_{\mathcal{C}} + b_{\mathcal{LN}}) = A^T b_{\mathcal{C}}$ implying that the left-nullspace component of b causing the inequality is no longer part of the equation. Also observe that x became x^* because now there is a unique solution x^* to the problem $A^T Ax^* = A^T b$. Since $A^T A$ is an invertible nxn matrix, it is possible to solve for the least-squares solution x^* to obtain $x^* = (A^T A)^{-1} A^T b$. Finally, multiply both sides of the equation by A to obtain the projection $b_{\mathcal{C}} = Ax^* = A(A^T A)^{-1} A^T b$ and the projection matrix is $P_{\mathcal{C}} = [A(A^T A)^{-1} A^T]$.
11. The problem $Ax = b$ where the columns of A are independent and $b \neq 0$ have a unique solution in the rowspace of A because A does not have a nullspace, i.e., independent columns indicate a trivial nullspace. If this matrix A is not square, then it is possible that $Ax \neq b$. Clearly, the least

squares solution $x^* = (A^TA)^{-1}A^Tb$ for this problem must also be in the rowspace of A (since A has independent columns.)

12. The problem $Ax = b$ where the columns of A are independent and $b \neq 0$ have a unique solution in the rowspace of A because A does not have a nullspace, i.e., independent columns indicate a trivial nullspace. If this matrix A is not square, then it is possible that $Ax \neq b$. Clearly, the least squares solution $x^* = (A^TA)^{-1}A^Tb$ for this problem must also be in the rowspace of A (since A has independent columns.)
13. All projection matrices are "idempotent": $PP = P$. The reason for this is that once a vector $b_{\mathcal{C}}$ is obtained by projecting into its column space, i.e., $b_{\mathcal{C}} = P_{\mathcal{C}}b$, multiplying by $P_{\mathcal{C}}$ again will not change the value of $b_{\mathcal{C}}$ since $b_{\mathcal{C}}$ already lies in the column space. Thus, $P_{\mathcal{C}}b_{\mathcal{C}} = P_{\mathcal{C}}b$ and substituting $b_{\mathcal{C}}$= $P_{\mathcal{C}}b$ into the left side of the equation yields $P_{\mathcal{C}}P_{\mathcal{C}}b_{\mathcal{C}} = P_{\mathcal{C}}b$ or $[P_{\mathcal{C}}P_{\mathcal{C}} - P_{\mathcal{C}}]b = 0$. Since this is true for any b, it follows that $[P_{\mathcal{C}}P_{\mathcal{C}} - P_{\mathcal{C}}] = [0]$ or $P_{\mathcal{C}}P_{\mathcal{C}} = P_{\mathcal{C}}$.
14. The converse is also true, all idempotent matrices are projection matrices onto their own column spaces.
15. The projection matrix $P_{\mathcal{LN}}$ that projects a vector v into the left-nullspace of A is obtained by realizing that v can be split into the sum of two orthogonal components. One component is obtained by projecting v onto the column space of A. It follows that the other orthogonal component must be the Left-nullspace component. Thus

    $$v_{\mathcal{LN}} = v - P_{\mathcal{C}}v = Iv - P_{\mathcal{C}}v = [I - P_{\mathcal{C}}]v$$

    where I is the identity matrix. It follows that $P_{\mathcal{LN}} = [I - P_{\mathcal{C}}]$.
16. To project into the rowspace it suffices to create a projection matrix onto the columnspace of $B = A^T$. Note that the columnspace of B is the rowspace of A. Thus, the projection matrix into the rowspace of A is: $P_{\mathcal{R}} = [B(B^TB)^{-1}B^T] = [A^T(AA^T)^{-1}A]$.
17. To obtain the projection onto the nullspace of A one follows the same logic used to obtain the projection onto the left-nullspace of A to obtain: $P_{\mathcal{N}} = [I - P_{\mathcal{R}}]$. All projection matrices project onto their own column space. Proof: P$\mathrm{b}_{\mathcal{C}} = \mathrm{b}_{\mathcal{C}}$ implies that every vector $\mathrm{b}_{\mathcal{C}}$ in the columnspace of A is also in the columnspace of P, i.e., $\mathrm{b}_{\mathcal{C}}$ is in the columnspace of P. Conversely, $\mathrm{Pb} = \mathrm{b}_{\mathcal{C}}$ implies that every vector Pb in the columnspace of P is equal to a vector $\mathrm{b}_{\mathcal{C}}$ in the columnspace of A. It follows that the columnspace of P is the same as the columnspace of A.
18. A projection matrix into a given subspace is unique. Proof: Assume P_1 and P_2 are nxn projection matrices with the same columnspace. Consider any arbitrary vector b in the co-domain of these projection matrices. Vector b can be expressed in terms of the columnspace and left-nullspace components: $b = b_C + b_{LN}$. It follows that $P_1b_{LN} = P_2b_{LN} = 0$ and $P_1b = P_2b = b_C$. This indicates that for every possible choice of b, $(P_1 - P_2) \cdot b = 0$. This is only possible if $P_1 = P_2$.
19. With the exception of the trivial projection, i.e., the identity matrix, projection matrices are not invertible because their nullspace is not empty.
20. Simple substitutions can be used to show that $P_{\mathcal{R}}, P_{\mathcal{LN}}$, and $P_{\mathcal{N}}$ are each idempotent matrices.

Exam Questions:

1. Prove that the rowspace of A have the same number of independent vectors as the columnspace of A.
2. If the columns of matrix A are independent of each other, provide a clear, thorough and comprehensive explanation proving that A^TA is a square, invertible matrix.

3. Suppose the nxm matrix A has independent columns but it is not a square matrix. What are the constraints on n and m for this to be true? Explain your answer. What is the nullspace of A for this case? Explain your answer.
4. Show that the solution to the least squares problem $Ax \neq b$ is $x = (A^T A)^{-1} A^T b$. What are the requirements on matrix A for this formula to work?
5. Assuming that the least squares solution to the problem $Ax \neq b$ can be expressed as $x = (A^T A)^{-1} A^T b$, is part of this solution on the nullspace? Explain your answer.
6. Consider the problem $Ax \neq b$ where A is a nxm matrix with n independent columns, then multiply by A^T to obtain $A^T A x^* = A^T b$. Provide a clear, thorough and comprehensive explanation on a) why the inequality went away and b) why x became x^*.
7. Derive the equation for a projection matrix onto the columnspace of a matrix A. You will receive no points for only writing down the projection matrix equation. All the points will be assigned to the derivation process.
8. Prove that any projection matrix must idempotent and explain why this should be intuitively so.
9. Prove that $P_{\mathcal{LN}}$, the projection matrix onto the left-nullspace of a matrix A must is indeed idempotent.
10. If f1(x) and f2(x) are two functions in the interval [1,-1], show how to use the projection matrix concept to project the function g(x) into the two dimensional subspace spanned by f1(x) and f2(x) using the typical definition of dot product for functions.
11. Given the projection P_c matrix for the columnspace of a matrix A. Derive (not just write the final formulas) projection matrices for the rowspace, leftnullspace, and nullspace of A. Clearly explain the logic process leading to your answers.
12. Prove that a projection matrix into a given subspace is unique.
13. Clearly explain why a matrix with a non-empty nullspace is not invertible.
14. Clearly explain why projection matrices are not invertible - with the exception of the trivial projection, i.e., the identity matrix.

Oct. 17 – Tuesday
LECTURE 16: Examples

Summary:

1. Example: triangulation. Suppose the distance of an object to n different known locations is given. The problem is to use this information to find the location of the object. The equation for the distance of the object with respect to location 1 can be mathematically expressed as

$$(p - p_1)^T (p - p_1) \cong d_1^2$$

where p is a vector containing the unknown coordinates of the location of the object, p_1 is a vector containing the known coordinates of location 1, and d_1 is the scalar distance between the object and location 1. The approximate, $\cong$, symbol is used because p_1 and d_1 are measured and, hence, subject to measurement errors. The equation above can be expanded to obtain,

$$p^T p - 2 p_1^T p + p_1^T p_1 \cong d_1^2$$

It is important to remember that $p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $p_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ are vectors containing the coordinates of the unknown position of the object, and the coordinates of the know location 1, respectively.

Using matrix notation, this equation can be expressed as

$$[-2p_1^T \quad 1] \begin{bmatrix} p \\ p^T p \end{bmatrix} \cong d_1^2 - p_1^T p_1$$

adding n more known distances and locations yield the following matrix problem

$$\underbrace{\begin{bmatrix} -2p_1^T & 1 \\ -2p_2^T & 1 \\ \vdots & \vdots \\ -2p_n^T & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} p \\ p^T p \end{bmatrix}}_{x} \cong \underbrace{\begin{bmatrix} d_1^2 - p_1^T p_1 \\ d_2^2 - p_2^T p_2 \\ \vdots \\ d_n^2 - p_n^T p_n \end{bmatrix}}_{b}$$

Therefore, it is seen that the triangulation problem can be seen as the problem

$$Ax \cong b$$

2. Example: consider the problem of fitting the quadratic polynomial $d = ax^2 + bx + c$ to four data points $\{(x_1, d_1), \cdots, (x_4, d_4)\}$ where the $d's$ are the data points subject to measurement errors. There is no expectation that a quadratic polynomial can perfectly fit all the data points, so using matrix notation, one can write

   $$\underbrace{\begin{bmatrix} {x_1}^2 & x_1 & 1 \\ {x_2}^2 & x_2 & 1 \\ {x_3}^2 & x_3 & 1 \\ {x_4}^2 & x_4 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{x} \cong \underbrace{\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}}_{b} \quad \text{or} \quad Ax \cong b$$

   The above problem can be solved in the least-squares sense by considering the orthogonal decomposition or expansion: $b = b_{\mathcal{LN}} + b_{\mathcal{C}}$. The least-squares idea is to make the equation exactly equal by stripping away $d_{\mathcal{LN}}$ to obtain $Ax = b_{\mathcal{C}}$. Still A is a 4x3 matrix and cannot be inverted to solve for x. This problem is circumvented by first multiplying both sides of the equation by A^T to obtain $A^T Ax = A^T b_{\mathcal{C}}$. Observe that

   $$A^T d = A^T(b_{\mathcal{LN}} + b_{\mathcal{C}}) = \overbrace{A^T b_{\mathcal{LN}}}^{0} + A^T b_{\mathcal{C}} = A^T b$$

   The matrix $[A^T A]$ is a 3x3 invertible matrix multiplying both sides of the last equation by $(A^T A)^{-1}$, one obtains the <u>least-squares solution</u>
   $x = (A^T A)^{-1} A^T b_{\mathcal{C}} = (A^T A)^{-1} A^T b$.
3. Observe that $d \cong ax^2 + bx + c$ is an expansion of vector d in terms of three non-orthogonal vectors x^2, x, and 1. The “least squares” process was used as a compromise to find the coefficients a, b, and c of the expansion when inexact information (data) about d was provided. In general, the process of finding the linear coefficients of an approximate expansion is called “<u>linear regression</u>”. The word “linear” is used because the coefficients are the weights of a linear combination of vectors.
4. In general, you can use linear regression to fit an equation of the form

$$d \cong a_1 f_1(x) + a_2 f_2(x) + \cdots + a_n f_n(x)$$

to obtain an equation of the form

$$\underbrace{\begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ \vdots & \vdots & \cdots & \vdots \\ f_1(x_m) & f_2(x_m) & \cdots & f_n(x_m) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}}_{x} \cong \underbrace{\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}}_{b}$$

or $Ax \cong b$.

Exam Questions:

1. Suppose the distance of an object to n different known locations is given. The problem is to use this information to find the location of the object. Show how to pose this problem such that it can be solved in the general matrix form $Ax \cong b$.
2. Show how to fit a second order polynomial in the variable x to a set of data points (x0,y0), (x1,y1),…, (x3,y3) using the concept of least squares. Assume the data points are corrupted with noise so a perfect fit is not possible.

Oct. 19 – Tuesday
LECTURE 17: Projection Matrices

Summary:

1. Recall from a previous lecture that four problems related to the matrix equation $Ax \cong b$:
   a) Invertible (perfectly constrained or statically determinate): $Ax = b$
   b) Over-constrained (Inconsistent or over-determined): $Ax \neq b$
   c) Under-constrained (under-determined, with infinite solutions): $Ax = b$
   d) Partly Under constrained and partly inconsistent: $Ax \neq b$

   a) is solved by inverting matrix A to obtain: $x = A^{-1}b$.
   b) can be addressed through the "least-squares" compromise: $Ax = b_{\mathcal{C}}$
   and $x^* = (A^T A)^{-1} A^T b$.
   c) can be addressed through the "shortest solution" optimization: $Ax_{\mathcal{R}} = b$.
   d) can be addressed using both ***"shortest solution" and "least squares"***: $Ax_{\mathcal{R}} = b_{\mathcal{C}}$.
2. Case (c) with infinite solutions, is often solved by finding the shortest of all possible solutions. This constitutes an optimization problem because the solution is a minimum length value. In practice, the shortest solution is readily found by forcing a solution $x_{\mathcal{R}}$ in the rowspace, i.e., do not allow a nullspace component in the solution. One approach for finding $x_{\mathcal{R}}$ is to first solve an intermediate matrix problem as follows. First, a basis for the rowspace of A is needed. One can perform a rref on A to find a basis of the rowspace of A - remember to select only the independent rows from the $rref(A)$ command. The basis vectors for the rowspace are then arranged as columns of a matrix R. Then $x_{\mathcal{R}}$ can be expressed as a linear combination of the columns of R, i.e.,

$$\mathrm{x}_{\mathcal{R}} = R^{\langle 0\rangle}z_0 + R^{\langle 1\rangle}z_1 + \cdots + R^{\langle r\rangle}z_r = Rz$$

Here r is the rank of the matrix, i.e., the number of linearly independent rows (or columns), and $z_0, z_1, \cdots, z_r$ are coefficients for the linear combination. Substituting $\mathrm{x}_{\mathcal{R}} = Rz$ into $A\mathrm{x}_{\mathcal{R}} = b$ one obtains $ARz = b$. Matrix (AR) can be shown to have independent columns by showing that there is no non-trivial linear combination of the columns of matrix AR that can add up to zero. Try it. To solve for z first multiply both sides of the equation by $(AR)^T$ to obtain $(AR)^T(AR)z = (AR)^Tb$ where $[(AR)^T(AR)]$ is an rxr square matrix that can be inverted to obtain $z = [(AR)^T(AR)]^{-1}(AR)^Tb$. The final solution is then $x_{\mathcal{R}} = Rz$ or $x_{\mathcal{R}} = R[(AR)^T(AR)]^{-1}(AR)^Tb$.

3. Note that since matrix (AR) have independent columns, it is not necessary to extract the linear independent columns of the original matrix A.
4. Surprisingly enough, the solution to (d) is obtained using the same final formula as for case (c). However, it is important to understand that both cases have a different philosophy. The shortest solution only yields, $ARz \neq b$. To remove the inequality, a least-squares compromise is needed. The compromise is to set $ARz = b_C$. The solution to this compromise is $x^* = R[(AR)^T(AR)]^{-1}(AR)^Tb$.
5. Case (d) is solved through the ***shortest least-squares solution***. The solution process based on using a rowspace solution together with a least squares compromise is also known as the ***Moore-Penrose Pseudo inverse***, i.e., $R[(AR)^T(AR)]^{-1}(AR)^T$ is the Moore-Penrose pseudo inverse. Later we will learn that this pseudoinverse can be much more conveniently obtained using the singular value decomposition.
6. It was shown in class how to obtain projection matrices onto the four spaces of a matrix using the QR decomposition. In particular, it was shown how to use the qr(.) function in Mathcad, along with the submatrix(), cols(), rows(), and rank() functions in Mathcad to obtain four matrices, each with orthogonal columns, representing an orthonormal basis for each of the four spaces of a matrix. It was also shown that, for a matrix Q with orthogonal columns, the matrix $P = QQ^T$represents a projection matrix onto the columnspace of Q.
7. Example of how derivatives of nonlinear equations can be used to generate linear constraint equations: finding the reaction force for a mass particle constrained to follow a circular motion. The position of the particle is given by a two-dimensional vector p. Newton's Law states that the $m \cdot p'' \ = \ f$, where p'' is the acceleration. The constraint equation is $p^Tp = R^2$, i.e., the particle follows a circular motion. The problem is to find the force f that must be exerted on the particle such that it follows a circular motion. Taking the derivative of the constraint we obtain $p'^T\ p + p^T\ p' = 0$. This is the same as $2p'^T\ p = 0$ or $p'^T\ p = 0$. Differentiating again, we obtain $p''^Tp + p'^Tp' = 0$. Multiplying by m: $m \cdot p''^Tp + m \cdot p'^Tp' = 0$. Substituting Newton's 2nd Law: $f^Tp + m \cdot p'^Tp' = 0$. Note that p'^Tp' is the same as v^2, i.e., the velocity squared of the particle. Let $p \ = \ u \cdot R$, i.e., a unit vector u representing the direction of p times the magnitude R of the position p. Thus $f^Tp = f^Tu \cdot R$. The final equation is then $f^Tu = -m \cdot \frac{v^2}{R}$, which states, that the component of the force in the radial direction, u, must be equal to the centripetal force. Observe that setting $A = f^T$, $x = p$, and $b = -m \cdot v^2$, this problem can be expressed as $Ax = b$ and the concepts of the four spaces of a matrix will apply.

Below are some facts about projection matrices that were not discussed in class, but you are expected to learn:

8. Proof that an idempotent matrix must be a projection matrix.
Let P be an $n \times n$ idempotent matrix and consider a vector b in the column space of P, i.e.,

$$b = b_c + b_{LN}$$

P will be a projection matrix if

$$Pb = P(b_c + b_{LN}) = \overbrace{Pb_c}^{b_c} + \overbrace{Pb_{LN}}^{0} = b_c$$

Therefore, we need to prove that $Pb_c = b_c$ and that $Pb_{LN} = 0$.

To prove that $Pb_c = b_c$ let $\{p_1, p_2, \cdots, p_r\}$ be a basis selected from the columns of P. Then b_c can be expressed as $b_c = \alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_r p_r$. However, since P is idempotent, i.e., $PP = P$, this implies that $Pp_j = p_j$ for $1 \le j \le r$. Therefore,

$$Pb_c = \; P(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_r p_r) = \alpha_1 \overbrace{Pp_1}^{p_1} + \alpha_2 \overbrace{Pp_2}^{p_2} + \cdots + \alpha_r \overbrace{Pp_r}^{p_r} = b_c$$

To prove that $Pb_{LN} = 0$ first observe that the equation $Pp_j = p_j$ implies p_j is both in the columnspace and in the rowspace of P. So it follows that $\{p_1, p_2, \cdots, p_r\}$ is also a basis for the rowspace. This also implies that every vector perpendicular to the columnspace will also be perpendicular to the rowspace, i.e., the nullspace and leftnullspace are equal: any vector in the leftnullspace will also be in the nullspace and will be converted to zero when multiplied by P, i.e., $Pb_{LN} = 0$.

9. Projection matrices are symmetric.
Proof:
Let A be a matrix containing independent columns from P, since P is unique, then it follows that $P = A(A^TA)^{-1}A^T$

$$P^T = [A(A^TA)^{-1}A^T]^T$$

Now we use the following matrix identity:

$$(\mathrm{PQ})^{\mathrm{T}} = \mathrm{Q}^{\mathrm{T}}\mathrm{P}^{\mathrm{T}} \qquad (1)$$

This identity can be used to show that $(\mathrm{PQR})^{\mathrm{T}} = \mathrm{R}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}}\mathrm{P}^{\mathrm{T}}$ as follows

$$(\mathrm{PQR})^{\mathrm{T}} = \mathrm{R}^{\mathrm{T}}(\mathrm{PQ})^{\mathrm{T}} = \mathrm{R}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}}\mathrm{P}^{\mathrm{T}} \qquad (2)$$

and to prove that that $(Q^T)^{-1} = (Q^{-1})^T$ as follows

$$I = (Q^{-1}Q)^T = (Q^T)(Q^{-1})^T$$

multiplying by $(Q^T)^{-1}$

$$(Q^T)^{-1} = (Q^{-1})^T \qquad (3)$$

Using (1), (2), and (3) above, it follows that

$$P^T = [A(A^TA)^{-1}A^T]^T = [A[(A^TA)^{-1}]^TA^T] = \; A(A^TA)^{-1}A^T = P$$

10. The rowspace and columnspace of a projection matrix are identical (spanned by the same vectors) because the projection matrices are symmetric, i.e., the rows of A are the same as the columns of A, so the spaces are spanned by the same vectors. Since projection matrices are square, this indicates that the nullspace and leftnullspace of projection matrices are identical as well.

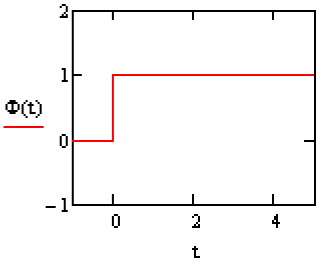
Exam Questions:

1. Explain why a matrix A with a non-empty nullspace is not invertible.
2. Consider the problem $Ax = b$ where A is not an invertible matrix. Explain what is the "shortest" solution to this problem and provide a logical explanation leading to a final equation for finding the "shortest" solution.
3. Given an $n \times m$ matrix A, and a matrix R containing a set of independent rows of A as its columns, prove that the matrix $B = AR$ has independent columns.
4. Explain all the detailed steps (not just list them, but provide the motivation and theoretical explanation behind each step) required to solve an under constrained linear (under-determined, with infinite solutions) problem of the form $Ax = b$.
5. Give a thorough logical explanation explaining how to obtain the shortest-least squares solution of the problem $Ax \neq b$.
6. Given a matrix Q with orthogonal columns, prove that the matrix $P = QQ^T$ represents a projection matrix onto the columnspace of Q.
7. Given that $(\mathrm{PQ})^{\mathrm{T}} = \mathrm{Q}^{\mathrm{T}}\mathrm{P}^{\mathrm{T}}$ prove that $(\mathrm{PQR})^{\mathrm{T}} = \mathrm{R}^{\mathrm{T}}\mathrm{Q}^{\mathrm{T}}\mathrm{P}^{\mathrm{T}}$ and that $(Q^T)^{-1} = (Q^{-1})^T$ provided Q is invertible.
8. Prove that projection matrices are symmetric.
9. Prove that the rowspace and the columnspace of projection matrices are identical.
10. Prove that the leftnullspace and the nullspace of projection matrices are identical.
11. Explain all the detailed steps (not just list them, but provide the motivation and theoretical explanation behind each step) required to solve an under constrained linear (under-determined, with infinite solutions) problem of the form $Ax = b$.
12. Prove that any idempotent matrix must be a projection matrix.

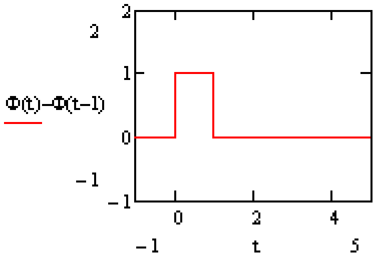
Oct. 24 – Tuesday
LECTURE 18: Pulse Response

Summary:

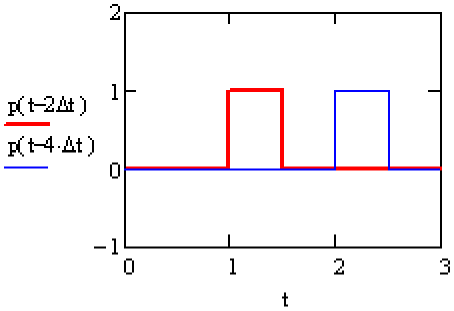
1. The Heaviside (unit) step function <u>can</u> be defined as: $\Phi(t) = \begin{cases} 1 & if\ \ t > 0 \\ 0 & otherwise \end{cases}$



2. If it is desired to start the pulse at time t_o instead of time zero, one only needs to shift the graph of the function to the right, i.e., use $\Phi(t-t_o)$.
3. Here is a link to Oliver Heaviside (http://en.wikipedia.org/wiki/Heaviside) short bio.
4. A pulse of unit height and width Δt is just the difference between two step functions: $p(t) = \Phi(t) - \Phi(t-\Delta t)$.



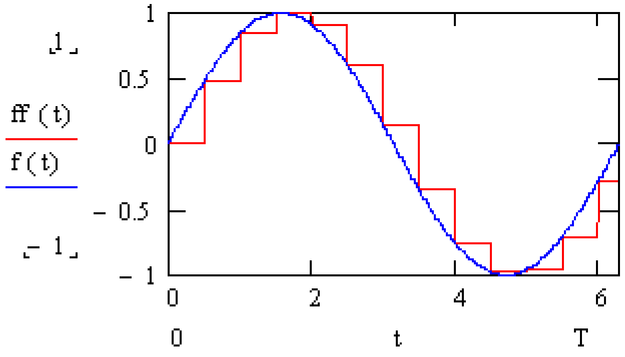
5. The pulse $p(t)$ can be translated or moved k time steps to the right by writing $p(t-k\Delta t)$. Observe that the pulses $p(t-k\Delta t)$ and $p(t-j\Delta t)$ do not overlap if $k \neq j$. This means that $p(t-k\Delta t)$ and $p(t-j\Delta t)$ are orthogonal to each other.



6. A function $f(t)$ can be approximated using a summation or "train" (think of a series of consecutive pulses as the wagons in a train) of orthogonal pulses (orthogonal expansion) as follows

$$ff(t) \cong \sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t)\, p(t-k\Delta t)$$

where T is the total length of time of interest and $\Delta t = T/N$.



7. For a linear differential equation with constant coefficients, the solution process is linear with respect to the input $f(t)$, and the solution can be expressed as $x(t) = L(f(t))$ where $L(\cdot)$ is a label given to the linear solution process. Since $f(t) \cong ff(t)$, then $x(t) \cong L\big(ff(t)\big)$ or

$$x(t) \cong L\left(\sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t)\, p(t - k\Delta t)\right)$$

Applying linearity,

$$x(t) \cong \sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t)\, L(p(t - k\Delta t))$$

Now, $f(k\Delta t)$ is readily available as an input to the differential equation and $L(p(t - k\Delta t))$ is the response (solution) of the differential equation to the pulse $p(t - k\Delta t)$.

8. Linear differential equations with constant coefficients (in contrast with time-varying coefficients) are time-invariant because the manner in which they respond to inputs is not changed (invariant) if the same inputs are applied at a different time: if a mass-spring-damper is kicked (simulating a pulse input) it will vibrate and eventually the vibrations will stop due to the damping; if then another kick is applied again in exactly the same manner, the response would be identical to the response the first time the mass-spring-damper was kicked, except that this latter response happens later in time. These kind of systems are named "LTI" for "Linear-Time-Invariant". Linear differential equations with constant coefficients are LTI.
9. Let the response to a pulse $p(t)$ of unit magnitude and width Δt be labeled $r(t) = L(p(t))$. Since in addition to linear, the system ($x'' + 2x' + 2x = f(t)$) is time-invariant, then $L(p(t - k\Delta t)) = r(t - k\Delta t)$. This means that

$$x(t) \cong \sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t)\, r(t - k\Delta t)$$

where $r(t)$ is the solution to the differential equation

$$r'' + 2r' + 2r = p(t)$$

10. The importance of the above is that if the pulse-response $r(t)$ is available, then the response of the "system" described by the differential equation to an arbitrary input $f(t)$ can be readily

approximated. The pulse-response $r(t)$ can be obtained directly by physically measuring the actual response to an input pulse. In this case there will be no need for an analytical solution. The pulse response $r(t)$ can also be found by applying a pulse $p(t)$ to a linear, finite element model (probably defined by a complex geometry or arrangement of beams, gears, etc.) of a system and then storing the response $r(t)$ at a location of interest.

11. The time constant of a LTI system can be used to find how long the response $r(t)$ is before decaying to zero. Usually the response $r(t)$ is assumed to be zero after 4.6 times the dominant (largest) time-constant of the system. For differential equations in space rather than in time, Saint Venant's principle provides the same idea: "... the difference between the effects of two different but statically equivalent loads becomes very small at sufficiently large distances from load." "The Saint-Venant's principle can be regarded as a statement on the asymptotic behavior of the Green's function by a point-load." [http://en.wikipedia.org/wiki/Saint-Venant%27s_Principle]
12. The width of a pulse used to approximate a function is usually selected to be ten times shorter than the smallest time constant of interest. The reasoning for this is that it takes a time of 4.6 time constants for the response to stabilize (settle) to constant input. This means that the effect of a pulse of 1/10 the smallest time constant will have more than enough time to settle within the fastest response rate of the system. That is, the fastest response rate is the response due to the smallest time constant: remember that partial fraction expansion separates the response of a linear differential equation with constant coefficients into the sum of responses of several partial fractions, each partial fraction associated with a time constant.
13. The pulse-response can also be found directly by solving the differential equation $r'' + 2r' + 2r = p(t)$. However, since $p(t) = \Phi(t) - \Phi(t - \Delta t)$, it is only necessary to find the solution to the differential equation $s'' + 2s' + 2s = \Phi(t)$ where $s(t)$ is the response (solution) of the differential equation to a step input. Using linearity, it follows that $r(t) = s(t) - s(t - \Delta t)$.
14. To obtain the step response, the Laplace Transform of the differential equation

$$s'' + 2s' + 2s = \Phi(t)$$

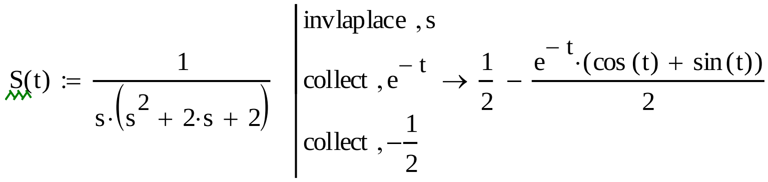
was applied to obtain

$$s^2 S + 2sS + 2S = 1/s$$

Solving for S:

$$S(s) = \frac{1}{s(s^2 + 2s + 2)}$$

Finally, $S(t)$ was obtained by applying the invlaplace symbolic function in Mathcad.



$$S(t) := \frac{1}{s\cdot\left(s^2 + 2\cdot s + 2\right)} \left| \begin{array}{l} \text{invlaplace}, s \\ \text{collect}, e^{-t} \\ \text{collect}, -\frac{1}{2} \end{array} \right. \rightarrow \frac{1}{2} - \frac{e^{-t}\cdot(\cos(t) + \sin(t))}{2}$$

15. It is very important to ensure that $S(t) = 0$ for $t < 0$, i.e., the unit step response should remain zero before the step is applied. This led to redefining the solution $S(t)$ in Mathcad so that $S(t) = 0$ for $t < 0$, i.e., multiplying the previous solution by $\Phi(t)$.

16. The summation

$$x(t) \cong \sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t)\, L(p(t - k\Delta t))$$

can be converted into an integral by first multiplying by (Δt/Δt) - this step is needed to generate the dt term in the integral.

$$x(t) \cong \sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t) \frac{L(p(t - k\Delta t))}{\Delta t} \Delta t$$

Taking the limit as $\Delta t \to 0$ leads to the desired integral.

17. More insight can be gained into the limit process by considering the limit of $L(p(t))/\Delta t$ as $\Delta t \to 0$ in more detail. First, let

$$\lim_{\Delta t \to 0} \frac{L(p(t))}{\Delta t} = \lim_{\Delta t \to 0} \frac{S(t) - S(t - \Delta t)}{\Delta t} = \frac{dS(t)}{dt} = \dot{S}(t)$$

18. The desired integral can now be obtained as follows.

$$x(t) = \lim_{\Delta t \to 0} \sum_{k=0}^{floor(t/\Delta t)} f(k\Delta t) \frac{L(p(t - k\Delta t))}{\Delta t} \Delta t$$

The summation becomes an integral as Δt becomes infinitesimally small while $N = T/\,\Delta t$ becomes infinitely large. Observe that $k\Delta t$ will vary from 0 to $T = N\Delta t$ as $\Delta t \to 0$. A new variable name, τ, with $0 \le \tau \le T$, is used to represent $k\Delta t$. Since Δt produces a change in $k\Delta t$, it follows that Δt translates into $d\tau$ in the integral:

$$x(t) = \int_0^t f(\tau)\dot{S}(t - \tau)d\tau$$

19. In order to simplify the notation, let

$$h(t) = \dot{S}(t)$$

be the impulse response of a LTI (Linear-Time-Invariant) differential equation. Then, the above integral can be expressed as

$$x(t) = \int_0^t f(\tau)h(t - \tau)d\tau$$

This integral is also known as the “Convolution Integral”.

20. Since the derivative of the step response is the impulse response, i.e.,

$$h(t) = \dot{S}(t)$$

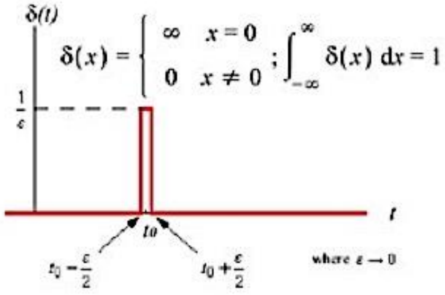
or

$$H = sS$$

and since

$$s^2S + 2sS + 2S = 1/s$$

21. Review of the definition Dirac delta function δ(t): as seen in the figure below, the area under the function remains equal to 1 as ε → 0. Also, as ε → 0, the height approaches infinity, while the width of the function approaches zero. Conceptually, it is easy to imagine the derivation of this function using a rectangle, but the shape is really not important because there are many other ways to achieve an area of 1 as the width of a function approaches zero.



The Laplace Transform of the Dirac delta function is equal to 1, i.e.,

$$\mathcal{L}(\delta(\mathrm{t})) = \int_{0^-}^{\infty} \mathrm{e}^{-\mathrm{s}\cdot\mathrm{t}}\delta(\mathrm{t})\mathrm{dt} = \int_{0^-}^{0^+} \underbrace{\mathrm{e}^{-\mathrm{s}\cdot\mathrm{t}}}_{\substack{\approx \mathrm{e}^{-0\cdot\mathrm{t}}=1 \\ \text{in such a} \\ \text{short time}}} \delta(\mathrm{t})\mathrm{dt} = \int_{0^-}^{0^+} \delta(\mathrm{t})\mathrm{dt} = 1$$

or $\mathcal{L}(\delta(\mathrm{t})) = 1.$

22. Based on the definition of the Laplace Transform of a Dirac delta function, it is seen that the equation $s^2H + 2sH + 2H = 1$ is the Laplace transform of the equation $h'' + 2h' + 2h = \delta(t)$. That is, the impulse response is the response to a Dirac delta “impulse” function.

Solving $s^2H + 2sH + 2H = 1$ for H yields,

$$H = \frac{1}{s^2 + 2s + 2}$$

which provides a very simple way to obtain the impulse response:

$$h(t) = \mathcal{L}^{-1}\left(\frac{1}{s^2 + 2s + 2}\right)$$

23. A shortcut notation for the convolution process has been defined using the “star” notation (*) as follows:

$$f * h = \int_0^t f(\tau)h(t-\tau)d\tau$$

Thus,

$$x(t) = \; f * h$$

24. The importance of the convolution integral is that the Laplace Transform of the Convolution integral is given as

$$\mathcal{L}(f * h) = \mathcal{L}\left(\int_0^t f(\tau)h(t-\tau)d\tau\right) = F(s) \cdot H(s)$$

That is, convolution in time is the same as multiplication in the Laplace domain.

25. Here is a link (http://en.wikipedia.org/wiki/Fundamental_solution) for applying convolution to Partial Differential Equations (PDEs).
26. Here is another link (http://en.wikipedia.org/wiki/Green%27s_function) relating Green's Functions for solving PDEs using convolution as discussed in class.

27. In order to simplify the notation let $h(t) = \dot{S}(t)$ be the impulse response of a LTI differential equation. Then, the "Convolution Integral" can be expressed as

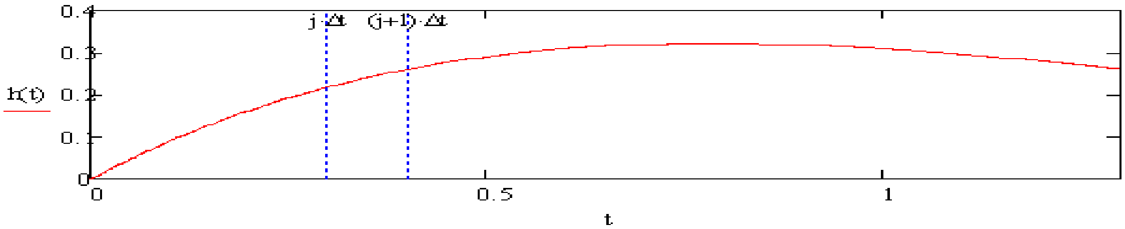
$$x(t) = f * h = \int_0^t f(\tau)h(t-\tau)d\tau$$

Substituting $t = k\Delta t$, the above integral can be split into several pieces as follows,

$$x(k\Delta t) = \int_0^{\Delta t} f(\tau)h(k\Delta t-\tau)d\tau + \int_{\Delta t}^{2\Delta t} f(\tau)h(k\Delta t-\tau)d\tau + \cdots + \int_{(k-1)\Delta t}^{k\Delta t} f(\tau)h(k\Delta t-\tau)d\tau$$

Inside each integral, the value $h(k\Delta t-\tau)$ varies with respect to τ. However, the figure below shows that Δt can be chosen small enough such that $h(\cdot)$ remains relatively unchanged in any interval of length Δt. To ensure this, the value of Δt should be chosen to be smaller than the smallest time constant of interest, for example, ten times smaller may be an adequate value for many applications. The upper limit of each piecewise integral can be used to obtain a constant to use in lieu of $h(k\Delta t-\tau)$ within each integral:

$$x(k\Delta t) = \int_0^{\Delta t} f(\tau)h(k\Delta t-\Delta t)d\tau + \int_{\Delta t}^{2\Delta t} f(\tau)h(k\Delta t-2\Delta t)d\tau + \cdots + \int_{(k-1)\Delta t}^{k\Delta t} f(\tau)h(k\Delta t-k\Delta t)d\tau$$



Moving the constant values of $h(\cdot)$ outside each integral:

$$x(k\Delta t) = h\big((k-1)\Delta t\big)\int_0^{\Delta t} f(\tau)d\tau + h\big((k-2)\Delta t\big)\int_{\Delta t}^{2\Delta t} f(\tau)d\tau + \cdots + h\big((k-k)\Delta t\big)\int_{(k-1)\Delta t}^{k\Delta t} f(\tau)d\tau$$

Defining $x_k \equiv x(k\Delta t)$ and $h_k \equiv h(k\Delta t)$ simplifies the above equation to

$$x_k = h_{k-1}\int_0^{\Delta t} f(\tau)d\tau + h_{k-2}\int_{\Delta t}^{2\Delta t} f(\tau)d\tau + \cdots + h_0\int_{(k-1)\Delta t}^{k\Delta t} f(\tau)d\tau$$

Defining $x_k \equiv x(k\Delta t)$ and $h_k \equiv h(k\Delta t)$ simplifies the above equation to

$$x_k = h_{k-1}\int_0^{\Delta t} f(\tau)d\tau + h_{k-2}\int_{\Delta t}^{2\Delta t} f(\tau)d\tau + \cdots + h_0\int_{(k-1)\Delta t}^{k\Delta t} f(\tau)d\tau$$

Furthermore, define $A_j = \int_{j\Delta t}^{(j+1)\Delta t} f(\tau)d\tau$ to obtain

$$x_k = h_{k-1}A_0 + h_{k-2}A_1 + \cdots + h_0A_{k-1} = \sum_{j=0}^{k-1} h_jA_{k-1-j}$$

The A_j's can be interpreted as areas under the input function $f(t)$.

The equation

$$x_k = \sum_{j=0}^{k-1} h_jA_{k-1-j}$$

is also known as the "Discrete Convolution" of h and A, i.e,

$$x = h * A$$

Where the * symbol results in a continuous convolution or a discrete convolution depends on whether the variables are discrete or continuous.

Exam Questions:

1. Write down the mathematical definition of the Heaviside's step function?
2. Show how to create a rectangular pulse function of unit height and width Δt, with the pulse triggering at time t=0. Then show how to create another pulse function of the same width and height, but triggering at time t_o.
3. Show how a function $f(t)$ can be approximated using a summation or "train" of orthogonal pulses. That is, provide a clear explanation on the logical process used to finally arrive at a formula for the approximation.
4. What is an LTI system and what kind of differential equations are used to model LTI systems?
5. Provide a clear logical explanation explaining how to use a set of orthogonal pulse functions and the definition of linearity to obtain an approximate solution to a differential equation used to model an LTI system.
6. Are there any guidelines for the width of the pulses used in question 5 above? Explain your answers.

7. Use linearity, the pulse response, and summation notation to derive the convolution integral as applied to the solution to a LTI differential equation. You must clearly explain the logical steps behind every detail of the process. During your derivation you must also show, using the limit process, that the impulse response is the derivative of the step response.
8. Show how to discretize the convolution integral as a summation of the products of areas under the input f(t) and discretized values of the impulse response, i.e.,

$$x_k = \sum_{j=0}^{k-1} h_j A_{k-1-j}$$

   As usual, you must clearly explain the logic behind each step of your derivation.
9. For the convolution integral, what value of Δt should be chosen such that

$$\int_0^t f(\tau)h(t-\tau)d\tau \cong\ h_{k-1}\int_0^{\Delta t} f(\tau)d\tau + h_{k-2}\int_{\Delta t}^{2\Delta t} f(\tau)d\tau + \cdots + h_0\int_{(k-1)\Delta t}^{k\Delta t} f(\tau)d\tau\ ?$$

   Provide a clear, logical explanation for your answer.

Oct. 26 – Thursday

LECTURE 19: Pulse Response - continuation

Summary:

1. Review of Lecture 18: detailed review of impulse function, impulse response, and convolution. A pdf file on "Convolution and PWM" has been placed on the Notes section of the course website.
2. From the previous lecture, the convolution integral can be approximated as

$$x_k = h_{k-1}\int_0^{\Delta t} f(\tau)d\tau + h_{k-2}\int_{\Delta t}^{2\Delta t} f(\tau)d\tau + \cdots + h_0\int_{(k-1)\Delta t}^{k\Delta t} f(\tau)d\tau$$

   where $x_k \equiv x(k\Delta t)$ and $h_k \equiv h(k\Delta t)$.

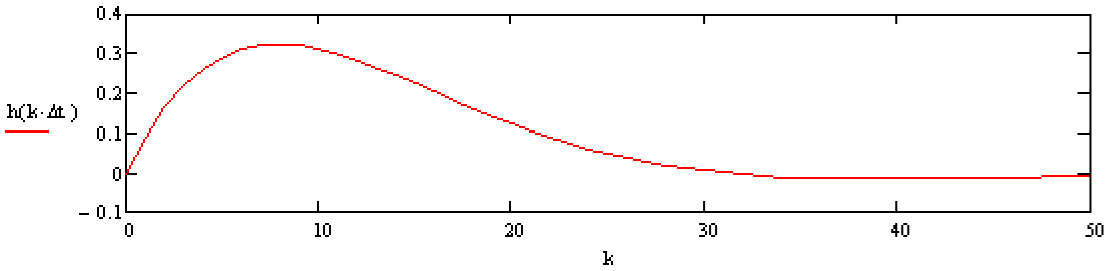
   Furthermore, defining $A_j = \int_{j\Delta t}^{(j+1)\Delta t} f(\tau)d\tau$ one obtains

$$x_k =\ h_{k-1}A_0 +\ h_{k-2}A_1 + \cdots +\ h_0 A_{k-1} = \sum_{j=0}^{k-1} h_j A_{k-1-j}$$

   The A_j's can be interpreted as areas under the input function $f(t)$.

   Observe that the indices of the h and A sequences in reverse order of each other. For many problems, the impulse response decays to zero for large values of k. For example, in the figure below, the impulse response decays to a value very close to zero after $k \geq N$, where $N = 40$.

Thus, for $k \geq N$, the summation can be expressed in two parts as follows.



$$x_k = \sum_{j=0}^{N-1} h_j A_{k-1-j} + \overbrace{\sum_{j=N}^{k-1} h_j A_{k-1-j}}^{\cong 0}$$

The second summation on the right is approximately zero because, as noted before, the values of $h_j \cong 0$ for $k \geq N$. The "efficient" calculations for finding x_k with a minimum number of calculations are summarized below.

$$x_k = \begin{cases} \displaystyle\sum_{j=0}^{k-1} h_j A_{k-1-j} & if\ 0 < k < N \\ \displaystyle\sum_{j=0}^{N-1} h_j A_{k-1-j} & if\ k \geq N \end{cases}$$

Note that the summation at the top has fewer terms than the summation at the bottom. Also, the summation at the bottom adds a fixed number N of terms while the number of terms added at the top summation varies with k. The two summations can be combined into a single summation by adding N zeros to the beginning of sequence A. The new longer sequence can be labeled B. Observe that $B_{k+N} = A_k$ and performing this substitution (wherever there is a k in A, there should be a $(k+N)$ in B) one obtains

$$x_k = \sum_{j=0}^{N-1} h_j B_{k+N-1-j} \quad \text{for} \quad k \geq 0$$

Careful observation shows that the latter summation performs the same calculations as the two summations previously defined. This latter summation goes by the name of "**Finite Impulse Response (FIR)** " because only N values of the impulse response are needed. Be careful to remember that the FIR works only when the impulse response decays to zero.

In Mathcad one can define $\mathrm{B}_{\mathrm{k+N}} := \mathrm{A}_{\mathrm{k}}$ for a specified range of values of k. Thus, $\mathrm{B}_{\mathrm{N}} = \mathrm{A}_0$, $\mathrm{B}_{\mathrm{N+1}} = \mathrm{A}_1$, etc. If this is the first time B is defined, then Mathcad will automatically assign values of zero to $\mathrm{B}_0, \mathrm{B}_1, \cdots, \mathrm{B}_{\mathrm{N-1}}$.

3. Two applications can be contemplated by considering the areas $A_j = \int_{j\Delta t}^{(j+1)\Delta t} f(\tau) d\tau$. The first application assumes that $f(t)$ remains approximately constant inside each piecewise integral.

$$A_j = \int_{j\Delta t}^{(j+1)\Delta t} \underbrace{f(\tau)}_{\cong f(j\Delta t)}\, d\tau \;\cong f(j\Delta t)\overbrace{\int_{j\Delta t}^{(j+1)\Delta t} d\tau}^{\Delta t} \;= f(j\Delta t)\Delta t$$

defining $f_j \equiv f(j\Delta t)$, the above equation reduces to $A_j = f_j\Delta t$ or $B_{k+N} = A_k = f_k\Delta t$.

Thus, the Finite Impulse Response can be expressed as:

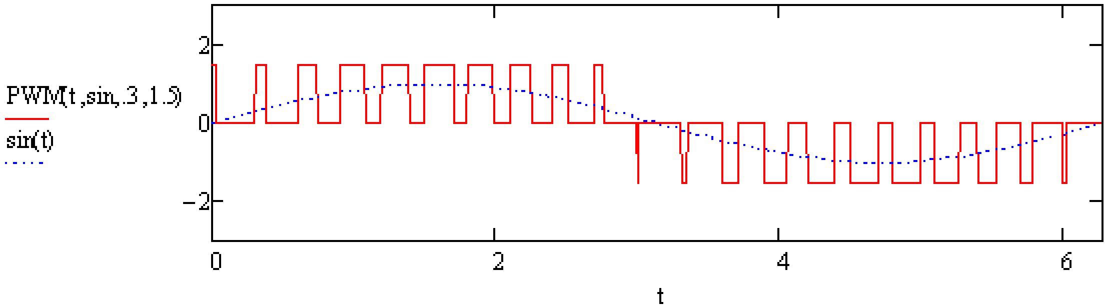
$$x_k = \begin{cases} \sum_{j=0}^{k-1}\left(h_j \cdot \Delta t \cdot f_{k-1-j}\right) & if\ k < N \\ \sum_{j=0}^{N-1}\left(h_j \cdot \Delta t \cdot f_{k-1-j}\right) & if\ k \geq N \end{cases}$$

As before in item 2, one can set $F_{k+N} = f_k$ to obtain

$$x_k = \sum_{j=0}^{N-1} h_j \cdot \Delta t \cdot F_{k+N-1-j} \quad \text{for} \quad k \geq 0$$

Again, recall that in Mathcad one can define $\mathrm{F}_{\mathrm{k+N}} := \mathrm{f}_{\mathrm{k}}$ for a specified range of values of k. Thus, $\mathrm{F}_{\mathrm{N}} = \mathrm{f}_0$, $\mathrm{F}_{\mathrm{N+1}} = \mathrm{f}_1$, etc. If this is the first time F is defined, then Mathcad will automatically assign values of zero to $\mathrm{F}_0, \mathrm{F}_1, \cdots, \mathrm{F}_{\mathrm{N-1}}$.

4. The second application takes advantage of the fact that only the areas $\mathrm{A}_{\mathrm{j}} = \int_{\mathrm{j}\Delta\mathrm{t}}^{(\mathrm{j}+1)\Delta\mathrm{t}} \mathrm{f}(\tau)\mathrm{d}\tau$ are needed to obtain the response x_{k}. Any other function that produces the same areas will yield the same response. One particular function of much practical interest consists of a sequence of rectangular pulses having the same height but varying widths. There is one pulse for each Δt interval, i.e., one pulse for each area A_{j}. The width of each pulse is varied so that the area under the j-th pulse matches the area of A_{j}. If the height common to all the pulses is denoted as H, the width of the j-th pulse is $\mathrm{w}_{\mathrm{j}} = \mathrm{A}_{\mathrm{j}}/\mathrm{H}$. This application is known as “Pulse Width Modulation (PWM).” The height H must be larger than the maximum value of f(t), this guarantees that $\mathrm{w}_{\mathrm{j}} < \Delta\mathrm{t}$. The figure below provides an example of PWM. The importance of PWM is that a simple device can be designed to create pulses with varying widths. In power electronics, power amplifiers can be designed to produce a PWM output. When applied to systems such as speakers, motors, and other devices, the PWM output of the power amplifier elicits the same dynamic response as the original, non PWM signal. The PWM signal is typically generated by power transistors that are switched ON or OFF to vary the width of the rectangular pulses. PWM amplifiers are popular because they are energy efficient in the sense that they do not generate nearly as much heat as regular analog amplifiers. The reason for this energy efficiency is that transistors have a very small resistance to current flow when operated fully ON.



Exam Questions:

1. Starting with the convolution integral and assuming that the impulse response of a given LTI differential equation approaches zero in an exponential fashion, provide a thorough derivation of how to obtain the Final Impulse Response equation.
2. Starting with the convolution integral, clearly explain the theoretical details justifying the process of of pulse width modulation (PWM). Your derivation must clearly show how to find the width of each pulse, assuming a constant height H for all pulses.
3. For a given continuous function $f(t)$ and a given Δt value for the equation $A_j = \int_{j\Delta t}^{(j+1)\Delta t} f(\tau)d\tau$, clearly explain what are the limitations on the height H of all pulses used in the PWM approximation. That is, can the height H be any numerical value? Explain your answer.

Oct. 31 - Tuesday
LECTURE 20: Complex Numbers

Summary:

1. Read all the notes on Convolution and/or Finite Impulse Response Filters placed on the Notes section of the course website.
2. Read the note on "Euler's Formula and Complex Numbers" placed on the Notes section of the course website.
3. A brief overview on the development of numbers. Initially, people could only count positive integers (1, 2, 3,....) perhaps due to the need to distribute goods among themselves. Introducing the number 0 was a significant abstract development that was probably difficult to accept by some people. Numbers can also be thought as arising from the need of people to solve equations. Equation-wise, integers were a result of trying to solve an equation such as x - 5 = 0. Equations such as x + 3 = 0 were impossible to solve until negative numbers were introduced. At that point in time negative numbers must have been very intriguing. Still, other equations such as 3x - 1= 0 could only be solved by introducing rational (fractional) numbers. The problem was complicated further by considering x^2-2=0. Thus, there was a need to add irrational numbers: these numbers never end, they have infinite decimal places in a non-repeating sequence! Still there were some equations that could not be solved. For instance x^2 + 1 = 0. Since irrational and rational numbers had been shown to completely fill a line (the real line), there was no space left to fit in more numbers. It turned out that the solution to x^2 + 1 = 0 belongs in another line (dimension). Hence, complex numbers are two-dimensional numbers, they have x and y components (real and imaginary). Because the roots of algebraic &

transcendental equations are complex numbers, the two dimensional space of complex numbers is the most appropriate numeric system to use when solving equations.

4. In Mathcad the imaginary number is entered as "1i" or "1j" (as a single expression, i.e., do not multiply the 1 and the i).
5. A complex number $x = a + jb$ can be interpreted as a vector: the real part correspond to the x-axis and the imaginary part corresponds to the y-axis component.
6. The magnitude of x is usually written as $|x|$ and is given by: $|x| = \sqrt{a^2 + b^2}$.
7. The angle or "phase angle" of a complex number $x = a + jb$ is given by $\theta = atan(b/a)$ - you may need to correct this formula to account for the correct quadrant where the complex number lies. For instance, the result of the equation $atan(1) = 45°$ may be due to $a = 1$ and $b = 1$ in which case the angle is correct, or it may be due to $a = -1$ and $b = -1$, resulting in $a/b = 1$ and $atan(a/b) = 45°$ but we know that $(-1, -1)$ lie in the third quadrant so the correct answer should be $180° + 45° = 225°$. The atan2(a,b) function in Mathcad provides the answer for the correct quadrant.
8. The equation $j^2 + 1 = 0$ is the characteristic polynomial of the differential equation $x'' + x = 0$. In order to solve the characteristic polynomial one would need to understand the meaning of j. This example relates the value of j to the solution of a physical problem: a mass-spring problem. Since the free solution of a mass spring damper is just a sustained vibration combining sine and cosine functions, one would suspect that the solution to the problem $j^2 + 1 = 0$ may be related to sine and cosine functions.
9. Euler's Formula ($e^{j\theta} = cos(\theta) + j\,sin(\theta)$) was obtained by comparing the Taylor series expansion of $sin(\theta)$, $cos(\theta)$, and e^{θ}, along with th equation $j^2 + 1 = 0$.

Exam Questions:

1. Why is it wrong to define the imaginary number i as $i \equiv \sqrt{-1}$? What would be the correct definition for i?
2. Explain how to find the phase angle and the magnitude of a complex number.
3. What is Euler's Formula.

Nov. 2- Thursday
LECTURE 21: Complex Numbers - continuation

Summary:

1. Let $x = a + jb$ be a complex number. The complex conjugate is denoted here as x^* and is given by $x^* = a - jb$ (reverse the sign of the imaginary part).
2. $x = a + jb$ can be expressed as $x = Re^{j\theta}$ where $R = \sqrt{a^2 + b^2}$ and $\theta = atan(b/a)$.
3. In general, the complex conjugate of any complicated expression can be obtained by reversing the sign of each imaginary term in the expression. For instance, if we express a complex number as $x = Re^{j\theta}$ then its complex conjugate is $x^* = Re^{-j\theta}$ (note the minus sign next to j.)
4. Observe that $x \cdot x^* = R^2$. Proof: $Re^{j\theta}Re^{-j\theta} = R^2 e^{j(\theta-\theta)} = R^2$.
5. Multiplying a complex number $R_1 e^{j\theta_1}$ by another complex number $R_2 e^{j\theta_2}$ will result in $R_1 R_2 e^{j(\theta_1+\theta_2)}$ which implies that the magnitude of $R_1 e^{j\theta_1}$ is "stretched" to a new magnitude $R_1 R_2$ and it is "rotated" forward by an amount θ_2 to a new phase angle $(\theta_1 + \theta_2)$. In particular, if $R_2 = 1$, then multiplying by $R_2 e^{j\theta_2}$ results in a rotating $R_1 e^{j\theta_1}$ by an amount θ_2 in the positive (forward) θ direction.

6. Similarly, dividing a complex number $R_1 e^{j\theta_1}$ by another complex number $R_2 e^{j\theta_2}$ will result in $(R_1/R_2)e^{j(\theta_1-\theta_2)}$ which implies that the magnitude of $R_1 e^{j\theta_1}$ is "scaled" to a new magnitude (R_1/R_2) and it is "rotated" backwards by an amount θ_2 to a new phase angle $(\theta_1 - \theta_2)$.
7. Other useful facts include: $x + x^* = 2a$ and $x - x^* = j2b$.
8. The above implies that or

$$a = \frac{x + x^*}{2} \quad \text{and} \quad b = \frac{x - x^*}{2j}$$

In particular, if $x = e^{j\theta} = \underbrace{cos(\theta)}_{a} + j \underbrace{sin(\theta)}_{b}$,

$$cos(\theta) = \frac{e^{j\theta} + e^{-j\theta}}{2} \quad \text{and} \quad sin(\theta) = \frac{e^{j\theta} - e^{-j\theta}}{2j}$$

9. Complex roots from polynomials with real coefficients always come in complex conjugate pairs.
10. A function is “analytical” at a point, if all the derivatives of the function exist at that point. A function f(x) will have a Taylor Series expansion about a given point x_0 only it the function is “analytical” at x_0. An “analytical function” is a function that is analytical everywhere, i.e., for every x.
11. Two important implications of Taylor Series that are important to remember are:
    1) A truncated Taylor Series results in a polynomial. Polynomials are well-behaved analytical functions that are relatively well understood.
    2) For analytical functions, a Taylor Series expansion shows that knowing all the derivatives at a given location provides enough information to reconstruct the entire function. That is, local information can be used to deduce more global behavior.
12. Taylor Series provide a convenient path to generalize functions that are initially developed using only real numbers. For instance, to generalize the sin(x) function to work with imaginary numbers one only needs to consider the Taylor Series of sin(x), (which would be a polynomial) and insert the complex value of x into the resulting polynomial. Recall that the Taylor Series expanded about zero goes by the name of “Maclaurin Series”.
13. When generalizing a mathematical operation, it is important to decide which properties take precedence during the generalization process. For instance, when generalizing the concept of inner (dot) product to include vectors with elements made up of complex numbers, the relationship between inner product and length of a vector takes precedence: the inner product of a vector with itself must yield the square of the length of the vector.
14. For vectors based on real numbers the inner product can be defined as $(x, y) = y^T x = x^T y$ for finite dimensional vectors and $(x, y) = \int_a^b x(t)y(t)dt$ for infinite dimensional vectors. Based on the concept behind Pythagoras Theorem, the square of the length of vector x is (x, x). However, if c is a vector is complex, i.e., made up of complex numbers, the previous definition of inner product will not result in $c^T c$ being the square of the length of c. This problem is fixed by first defining the length of a complex vector c as follows:

$$\|c\| \equiv \sqrt{\bar{c}^T c}$$

Then the definition of the inner product can be generalized to include complex numbers as follows:

$$(x, y) \equiv \bar{y}^T x \text{ for finite dimensional vectors}$$

or

$(x, y) = \int_a^b \bar{y}(t)x(t)dt$ for infinite dimensional vectors

where $\bar{y}$ is the complex conjugate of vector y. Using the generalized version of the inner product we have that

$$\|c\| \equiv \sqrt{(c, c)}$$

Since the complex conjugate operation on a vector of real numbers leaves the vector unchanged, this generalization works well, i.e., it works for real vectors and for complex vectors.

15. Define the Hermitian operation of a vector y as $y^H \equiv \bar{y}^T$, i.e., Hermitian = "complex conjugate transpose". Then the inner product for finite dimensional vectors based on complex numbers can be defined as $(x, y) \equiv y^H x$. The Hermitian operation also works on matrices: $A^H \equiv \bar{A}^T$.
16. Be careful to understand that for complex vectors, the inner product (x, y) is not necessarily equal to (y, x): the inner product for complex numbers is not commutative. That is, $y^H x$ is not necessarily equal to $x^H y$ unless both x and y are vectors are real. However it is always true that $(x, y) = \overline{(y, x)}$ or $y^H x = \overline{x^H y}$, i.e., (x, y) is the complex conjugate of (y, x).
17. In Mathcad, the dot product notation works for complex nx1 vectors with n>1: $x \cdot y \equiv y^H x$. However, in Mathcad the dot product notation does not work when using scalars: $x \cdot y \neq \bar{y} \cdot x$ when y is a complex scalar.
18. The Gram-Schmidt orthogonalization process, as well as the projection matrix concept, are easily generalized to include complex vectors by using the inner product based on the Hermitian operation. For finite dimensional vectors this means replacing the transpose operation by the Hermitian operation (sometimes called Hermitian transpose
19. Review: The Fourier Series expansion is only defined for functions that are periodic. For a function $f(t)$ of period T, the Fourier Series expansion is

$$f(t) = C + \sum_{n=1}^{\infty} \left(A_n \sin(n\omega_0 t) + B_n \cos(n\omega_0 t)\right) \quad \text{where} \quad \omega_0 = \frac{2\pi}{T}$$

Since the $\sin(n\omega_0 t)$ and $\cos(n\omega_0 t)$ are orthogonal in the interval $[0, T)$, it follows that

$$A_n = \frac{\int_0^T f(t)\sin(n\omega_0 t)\,dt}{\int_0^T \sin(n\omega_0 t)^2\,dt} \quad B_n = \frac{\int_0^T f(t)\cos(n\omega_0 t)dt}{\int_0^T \cos(n\omega_0 t)^2 dt} \quad \text{and} \quad C = \frac{\int_0^T f(t)1dt}{\int_0^T 1^2 dt}$$

20. Review: Since $\int_0^{\mathrm{T}} \sin(\mathrm{n}\omega_0\mathrm{t})^2\,\mathrm{dt} = \int_0^{\mathrm{T}} \cos(\mathrm{n}\omega_0\mathrm{t})^2\mathrm{dt} = \frac{\mathrm{T}}{2}$, it follows that for $n = 1,2, \ldots$

$$A_n = \frac{2}{T}\int_0^T f(t)\sin(n\omega_0 t)\,dt, \quad B_n = \frac{2}{T}\int_0^T f(t)\cos(n\omega_0 t)dt, \quad \text{and} \quad C = \frac{1}{T}\int_0^T f(t)dt$$

Normalizing the $\sin(n\omega_0 t)$, $\cos(n\omega_0 t)$, and the 1 vectors, the Fourier Series expansion can be expressed as

$$f(t) = (C\sqrt{T}) \underbrace{\left(\frac{1}{\sqrt{T}}\right)}_{normalized} + \sum_{n=1}^{\infty}\left(\left(A_n\sqrt{\frac{T}{2}}\right)\underbrace{\left(\frac{\sin(n\omega_0 t)}{\sqrt{\frac{T}{2}}}\right)}_{normalized} + \left(B_n\sqrt{\frac{T}{2}}\right)\underbrace{\left(\frac{\cos(n\omega_0 t)}{\sqrt{\frac{T}{2}}}\right)}_{normalized}\right)$$

Observe that T is the square of the length of the "1" vector, and T/2 is the square of the length of the sine and the cosine vectors in the interval $[0, T)$.

21. In general, the terminology "mean value" is given to the "most representative value" of a variable which results in a given process applied to the variable and applied to the mean value remaining the same. For instance, consider the process resulting in the area under a function $f(t)$ in the interval $[0, T)$. The mean value in this case is

$$C = \frac{1}{T}\int_0^T f(t)\,dt$$

because

$$Area = \int_0^T f(t)\,dt = \int_0^T C\,dt = C \cdot T$$

Other mean values can be seen in the Notes section of the course website.

22. Pythagoras Theorem states that the square of the length of a vector is the sum of the square of the orthogonal components. For infinite dimensional vectors such as $f(t)$ in the interval $0 \leq t < T$ this is known as Parseval's Theorem. Since the sines and cosines used in the Fourier Series are orthogonal, then the square of the length of $f(t)$ in the interval $0 \leq t < T$ is given by

$$\int_0^T f(t)^2 dt = T \cdot C^2 + \frac{\mathrm{T}}{2} \cdot \sum_{n=1}^{\infty} [A_n^2 + B_n^2]$$

23. Here is a link (http://en.wikipedia.org/wiki/Fourier_series) on Fourier Series with an example on Fourier's motivation.

Exam Questions:

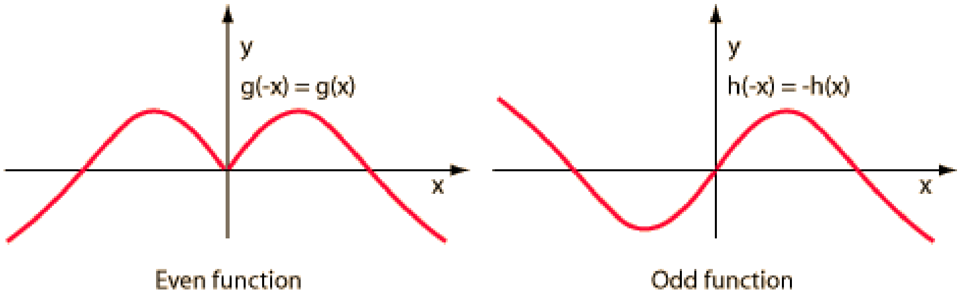
1. What are the two important implications of Taylor Series that were discussed in class and in the lecture notes?
2. For analytical functions based on real numbers, what would be a natural way to generalize these functions so that they can work with complex numbers as the arguments? Explain your answer. Why do the functions need to be analytical?
3. Is the inner product for vectors using complex numbers commutative? Explain your answer and provide an example.
4. How is the inner product defined when using vectors with complex numbers? What is the motivation behind this generalized definition of inner product? Clearly explain your answer.
5. Is the inner product for vectors using complex numbers commutative? Explain your answer and provide an example.
6. What is the Hermitian operator? How does it works?
7. For analytical functions based on real numbers, what would be a natural way to generalize these functions so that they can work with complex numbers as the arguments? Explain your answer. Why do the functions need to be analytical?
8. For what kind of functions is a Fourier Series defined? Write down the formula for the Fourier Series expansion (the "real" expansion, not the "complex" expansion) of a function that can be expanded using Fourier Series.

9. In the Fourier Series expansion of a real function, show how to find the length of the constant term, the sine terms, and the cosine terms. Then write down Parseval's Theorem and explain why this theorem works.

Nov. 7- Tuesday
LECTURE 22: Complex Fourier Series, Discrete Fourier Series

Summary:

1. In many applications engineers are interested in how a function behaves in a fixed window of time of width T. In that case, there is no harm in assuming the function is periodic with period T. In that case the function is said to be "extended periodically." This allows the use of Fourier Series and Parseval's Theorem as tools to analyze the function.
2. A function $f(t)$ is even if $f(t) = f(-t)$, i.e., the function is symmetric about the y-axis. The image of the function on the left side of the y-axis is the reflection of the function on the right side of the y-axis. If one plots the graph of $f(t)$ for $t \geq 0$ on graph paper and then fold the page with the crease exactly along the y-axis, if the ink is still wet, the reflection of the graph will be transferred to the opposite side of the y-axis. The resulting function will be an even function. The cosine function is an even function.
   A function $f(t)$ is odd if $f(-t) = -f(t)$, i.e., the function is "anti" symmetric about the y-axis: the image of the function on the left side of the y-axis is the "upside down" reflection of the function on the right side of the y-axis. The sine function is an odd function.



3. Any function can be expressed in terms of even and odd parts:

$$f(t) = f_{even}(t) + f_{odd}(t) \quad \text{where} \quad f_{even}(t) = \frac{f(t) + f(-t)}{2} \quad \text{and} \quad f_{odd}(t) = \frac{f(t) - f(-t)}{2}$$

   It follows that for any periodical function $f(t)$ of period T

$$f_{even}(t) = \sum_{n=0}^{\infty} \left(B_n \cos(n\omega_0 t)\right) \quad \text{and} \quad f_{odd}(t) = \sum_{n=1}^{\infty} (A_n \sin(n\omega_0 t))$$

   Also observe that $C = B_0$.

4. For real, periodic functions, the Fourier Series can be defined as

$$f(t) = C + \sum_{n=1}^{\infty}\left(A_n \sin(n\omega_0 t) + B_n \cos(n\omega_0 t)\right) \quad \text{where } \omega_o = \frac{2\pi}{T}$$

5. Here is a link (http://en.wikipedia.org/wiki/Fourier_series) on Fourier Series with an example on Fourier's motivation.

6. Fourier Series can be used to analyze an evenly and periodically extended function by using only even functions. This is done by defining a new evenly and periodically extended function $g(t)$ of twice the length (period) of $f(t)$ as follows:

   $$g(t) = \begin{cases} f(t) & if \quad T > t \geq 0 \\ f(-t) & if \quad -T \leq t < 0 \end{cases}$$

   Here $g(t)$ is an even function and assumed to be periodical with period $2T$ and

   $$g(t) = \sum_{n=0}^{\infty}\left(B_n \cos(n\omega_0' t)\right) \quad \text{where} \quad \omega_0' = \frac{2\pi}{2T} = \frac{\pi}{T} \quad \text{and} \quad B_n = \frac{\int_{-T}^{T} g(t)\cos(n\omega_0' t)dt}{\int_{-T}^{T} \cos(n\omega_0' t)^2 dt}$$

   Observe that since $g(t)$ and $\cos(n\omega_0' t)$ are both even functions, it follows that

   $$\int_{-T}^{T} g(t)\cos(n\omega_0' t)dt = 2\int_{0}^{T} f(t)\cos(n\omega_0' t)dt$$

   and

   $$\int_{-T}^{T} \cos(n\omega_0' t)^2 dt = 2\int_{0}^{T} \cos(n\omega_0' t)^2 dt$$

   so it follows that

   $$f(t) = \sum_{n=0}^{\infty}\left(B_n \cos(n\omega_0' t)\right) \quad for \quad T > t \geq 0$$

   where

   $$B_n = \frac{\int_{0}^{T} f(t)\cos(n\omega_0' t)dt}{\int_{0}^{T} \cos(n\omega_0' t)^2 dt} = \frac{2}{T}\int_{0}^{T} f(t)\cos(n\omega_0' t)dt \quad \text{for } n = 1,2,\cdots$$

   and

   $$B_0 = \frac{1}{T}\int_{0}^{T} f(t)dt$$

7. The Discretized version of the Cosine Series have many applications including the JPEG format used in digital pictures.
8. Similarly, it is possible to define a Sine Series using only odd functions and doubling the period of the original function.
9. Here is a link on Fourier Series with an example on Fourier's motivation.
10. Here is a link explaining Gibb's effect discussed in class.

11. A similar effect to Gibb's phenomenon occurs at the edges of functions that are approximated with polynomials. In this case the effect goes by the name of Runge's phenomenon.
12. Using complex notation, the above equation can be expressed as

$$f(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_o t} \quad \text{where} \quad \omega_o = \frac{2\pi}{T}$$

In order to see the relation between the real and the complex Fourier series, the latter series can be expressed as

$$f(t) = C_0 + \sum_{n=1}^{\infty} \left(C_n e^{jn\omega_o t} + C_{-n} e^{-jn\omega_o t} \right)$$

Using Euler's Formula, $e^{j\theta} = cos(\theta t) + jsin(\theta t)$, the above series becomes

$$f(t) = C_0 + \sum_{n=1}^{\infty} \left[\left(C_n cos(n\omega_o t) + C_{-n} cos(-n\omega_o t) \right) + j\left(C_n sin(n\omega_o t) + C_{-n} sin(-n\omega_o t) \right) \right]$$

Simplifying,

$$f(t) = C_0 + \sum_{n=1}^{\infty} [(C_n + C_{-n}) cos(n\omega_o t) + j(C_n - C_{-n}) sin(n\omega_o t)]$$

Comparing to the original Fourier Series from the previous bullet, we have that,

$$C = C_0, \quad B_n = (C_n + C_{-n}) \quad \text{and} \quad A_n = j(C_n - C_{-n})$$

Since B_n, and A_n are real constants, it follows $(C_n + C_{-n})$ and $j(C_n - C_{-n})$ must be real. This can only happen if C_n and C_{-n} are complex conjugates of each other.

Thus, it is seen that the complex Fourier Series is a generalization of the real Fourier Series, i.e., for the special case when $C_{-n} = \overline{C_n}$ (the bar on top means complex conjugate), the complex Fourier Series becomes identical to the real Fourier Series.

13. Recall that the complex Fourier Series can be expressed as

$$f(t) = \sum_{n=-\infty}^{\infty} C_k e^{jn\omega_o t} \quad \text{where} \quad \omega_o = \frac{2\pi}{T}$$

where the functions $e^{jn\omega_o t}$, for $n = 0,1, \cdots \infty$, are orthogonal in the interval $0 \leq \mathrm{t} < T$. That is, for integers n and m,

$$\left(e^{jn\omega_o t}, e^{jm\omega_o t} \right) = \int_0^T e^{jn\omega_o t} \cdot \overline{e^{jm\omega_o t}} dt = \int_0^T e^{jn\omega_o t} \cdot e^{-jm\omega_o t} dt$$

$$= \int_0^T e^{j(n-m)\omega_o t} dt = \begin{cases} 0 & \text{if } n \neq m \\ T & \text{if } n = m \end{cases}$$

where the 0 if $n \neq m$ is due to the fact that the function becomes a sum of a one or several integer periods of sine and cosine functions (the integral of a sine or cosine for a complete period is zero). On the other hand, if $n = m$, then $e^{j(n-m)\omega_o t} = e^0 = 1$ and $\int_0^T 1dt = T$.

Using the above properties for $e^{jn\omega_o t}$, the coefficients C_k can then be obtained using dot products for complex vectors as follows

$$C_n = \frac{(f(t), e^{jn\omega_o t})}{(e^{jn\omega_o t}, e^{jn\omega_o t})} = \frac{1}{T}\int_0^T f(t) \cdot e^{-jn\omega_o t} dt$$

14. The Complex Fourier Series pair can now be considered as a backward and forward process. In the forward process, the coefficients of the series are found. A frequency analysis examines the function by looking at the Fourier series coefficients and time is not considered. The backwards process reconstitutes the function in time using the coefficients.

$C_n = \frac{1}{T}\int_0^T f(t) \cdot e^{-jn\omega_o t} dt$ (Forward transformation)

and

$f(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_o t}$ (backwards)

A variation of Complex Fourier Series pairs include

$C_n = \frac{1}{T}\int_0^T f(t) \cdot e^{jn\omega_o t} dt$ and $f(t) = \sum_{n=-\infty}^{\infty} C_n e^{-jn\omega_o t}$

Observe that the sign of the exponentials have been swapped. Another variation is,

$C_n = \int_0^T f(t) \cdot e^{-jn\omega_o t} dt \;\; and \;\; f(t) = \frac{1}{T}\sum_{n=-\infty}^{\infty} C_n e^{jn\omega_o t}$

In this case observe that the normalization factor $\frac{1}{T}$ has been moved. Yet another variation is,

$C_n = \frac{1}{\sqrt{T}}\int_0^T f(t) \cdot e^{-jn\omega_o t} dt$ and $f(t) = \frac{1}{\sqrt{T}}\sum_{n=-\infty}^{\infty} C_n e^{jn\omega_o t}$

where the normalization factors are such that unit vectors are used both, in the forward and backwards formulae.

15. The forward formula $C_n = \frac{1}{T}\int_0^T f(t) \cdot e^{-jn\omega_o t} dt$ can be discretized by assuming that N data points are collected at uniform sampling intervals. Thus, the time between consecutive samples is $\Delta t = T/N$ and the integral can be approximated as

$$C_n = \frac{1}{T}\int_0^T f(t) \cdot e^{-jn\omega_o t} dt = \frac{1}{N\Delta t}\sum_{k=0}^{N-1} f(k\Delta t) e^{-jn\left(\frac{2\pi}{N\Delta t}\right)k\Delta t}\,\Delta t$$

or simplifying,

$$C_n = \frac{1}{N}\sum_{k=0}^{N-1} f_k e^{-j\omega_n k} \quad where\ \ \omega_n = n\left(\frac{2\pi}{N}\right)$$

Observe that C_n is just the projection of the vector f along the complex exponential vector. Since vector f is an n-dimensional vector, there would be n components, mainly, $C_0, C_1, \cdots, C_n$. Also observe that the square of the magnitude of the exponential vector is

$$\left(e^{-j\omega_n k}, e^{-j\omega_n k}\right) = \sum_{k=0}^{N-1} e^{-j\omega_n k}\, \overline{e^{-j\omega_n k}} = \sum_{k=0}^{N-1} e^{-j\omega_n k}\, e^{+j\omega_n k} = \sum_{k=0}^{N-1} e^0 = N$$

Thus, N is the square of the length of the exponential vector.

The original N data points in vector f can be recovered by summing up the orthogonal components:

$$f_k = \sum_{n=0}^{N-1} C_n e^{j\omega_n k} \quad where\ \ \omega_n = n\left(\frac{2\pi}{N}\right)$$

Below are several choices for forward and backward formulae depending on how the normalization factors are distributed:

a) $C_n = \frac{1}{N}\sum_{k=0}^{N-1} f_k e^{-j\omega_n k}$ forward and $f_k = \sum_{n=0}^{N-1} C_n e^{j\omega_n k}$ backward
b) $C_n = \sum_{k=0}^{N-1} f_k e^{-j\omega_n k}$ forward and $f_k = \frac{1}{N}\sum_{n=0}^{N-1} C_n e^{j\omega_n k}$ backward
c) $C_n = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} f_k e^{-j\omega_n k}$ forward and $f_k = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} C_n e^{j\omega_n k}$ backward

16. The Discrete Fourier Transform (DFT) of a N dimensional vector f is defined as choice (b) above

$$C_n = DFT(f)_n = \sum_{k=0}^{N-1} f_k e^{-j\omega_n k} \quad \text{where} \quad n = 0,1,\ldots,(N-1)\ \ and\ \ \omega_n = n\left(\frac{2\pi}{N}\right)$$

The Inverse Discrete Fourier Transform (IDFT) is the corresponding backward formula.

$$f_n = IDFT(C)_n = \frac{1}{N}\sum_{k=0}^{N-1} C_k e^{j\omega_n k} \quad \text{where} \quad n = 0,1,\ldots,(N-1)\ \ and\ \ \omega_n = n\left(\frac{2\pi}{N}\right)$$

Recall that $f(t)$ is sampled every Δt seconds to obtain N samples.

17. Note that the main purpose of $DFT(f)$ is to find the vector of coefficients C. The coefficients C are known as the frequency components of f.
18. Consider the angle of the complex exponential used in the DFT, $e^{-j\omega_n k}$. The angle is

$$-\omega_n k = -n\left(\frac{2\pi}{N}\right)k$$

Since adding or subtracting an integer multiple of 2π to the angle does not change the value of the complex exponential term

$$-n\left(\frac{2\pi}{N}\right)k - 2\pi k = -n\left(\frac{2\pi}{N}\right)k - 2\pi k\left(\frac{N}{N}\right) = -\frac{2\pi(N+n)k}{N} = -\omega_{N+n}k$$

It follows that

$$e^{-j\omega_n k} = e^{-j\omega_{N+n}k}$$

implying in turn, that $\quad C_n = C_{N+n}$

or, reversing the sign of n, $\quad C_{-n} = C_{N-n}$

and for real functions $\quad C_{-n} = \overline{C_n}$

so it follows that $\quad \overline{C_n} = C_{N-n}$

substituting $n = \frac{N}{2}$ yields $\quad \overline{C_{N/2}} = C_{N/2}$

The equations $C_{-n} = \overline{C_n}$, and $\overline{C_{N/2}} = C_{N/2}$ indicate that C_0 and $C_{N/2}$ must be real numbers because only a real number is equal to its complex conjugate.

In conclusion, the coefficients C_n are complex-conjugate symmetric about the index $N/2$. For example, using the equation $\overline{C_n} = C_{N-n}$ with $N = 6$, results in $\overline{C_0} = C_6,\ \overline{C_1} = C_5,\ \overline{C_2} = C_4, and\ \overline{C_3} = C_3$. For real functions $\overline{C_0} = C_6$ would also be real. The sequence of DFT coefficients can be seen in the table below. The first row of the table shows the first twelve DFT coefficients. The second row shows the equivalent coefficients.

C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
C_0	C_1	C_2	C_3	$\overline{C_2}$	$\overline{C_1}$	C_0	C_1	C_2	C_3	$\overline{C_2}$	$\overline{C_1}$

The above shows how the DFT of a function results in a repeating series of $N = 6$ coefficients because $C_n = C_{N+n}$. Moreover, $\overline{C_n} = C_{N-n}$, indicates that one only needs to know the coefficients from C_0 through $C_{N/2}$.

19. Not discussed in class: Given a vector f_{data} with N elements obtained by sampling a function $f(t)$ every Δt seconds, e.g., when collecting data using a data acquisition system, it is possible to find a continuous function $\hat{f}(t)$ interpolating all the data points by writing:

$$\hat{f}(t) = \frac{1}{N}\sum_{k=0}^{N-1} C_k e^{jk\omega_o t}$$

where the coefficients C_k are obtained using $DFT(f)$.
The number of calculations for the expression above can be reduced in half as follows.

$$\sum_{k=0}^{N-1} C_k e^{jk\omega_o t} = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} C_k e^{jk\omega_o t} = C_0 + \sum_{k=1}^{\frac{N-1}{2}} \left(C_k e^{jk\omega_o t} + C_{-k} e^{-jk\omega_o t}\right)$$

recalling that $C_{-k} = \overline{C_k}$ it is seen that the $C_{-k} e^{-jk\omega_o t}$ is the complex conjugate of $C_k e^{jk\omega_o t}$. Recalling that adding a number to its complex conjugate results in twice the real part results in:

$$\sum_{k=0}^{N-1} C_k e^{jk\omega_o t} = C_0 + \sum_{k=1}^{\frac{N-1}{2}} 2Re\left(C_k e^{jk\omega_o t}\right) = C_0 + 2Re\left(\sum_{k=1}^{\frac{N-1}{2}} \left(C_k e^{jk\omega_o t}\right)\right)$$

The advantage of the equation above is that there are only $\frac{N-1}{2}$ terms to add together. The final interpolation formula is then

$$f_{interp}(t) = \frac{1}{N}\left(C_0 + 2Re\left(\sum_{k=1}^{\frac{N-1}{2}} \left(C_k e^{jk\omega_o t}\right)\right)\right)$$

An example pdf file "Interpolating Data Using the DFT" has been placed in the files & notes section of the course website.

20. Not discussed in class: the following steps show how to express $DFT(f)$ as a matrix multiplication. First write $DFT(f)$ in long form:

$$\mathrm{C_n = DFT(f)_n} = \sum_{\mathrm{k=0}}^{\mathrm{N-1}} \mathrm{f_k e^{-jn\frac{2\pi}{N}k} = 1f_0 + e^{-jn\frac{2\pi}{N}1} f_1 + e^{-jn\frac{2\pi}{N}2} f_2 + \cdots + e^{-jn\frac{2\pi}{N}(N-1)} f_{N-1}}$$

Defining $\omega = e^{-j\frac{2\pi}{N}}$, C_n can be expressed as $C_n = 1f_0 + \omega^n f_1 + \omega^{2n} f_2 + \cdots + \omega^{(N-1)n} f_{N-1}$. The resulting matrix equation for the case of $N = 5$, is then

$$\begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & \omega^1 & \omega^2 & \omega^3 & \omega^4 \\ 1 & \omega^2 & \omega^4 & \omega^6 & \omega^8 \\ 1 & \omega^3 & \omega^6 & \omega^9 & \omega^{12} \\ 1 & \omega^4 & \omega^8 & \omega^{12} & \omega^{16} \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} \quad \text{where} \quad \omega = e^{-j\frac{2\pi}{N}}$$

The symmetric matrix with the ω's is called the Fourier matrix: "Fast Fourier Transform algorithms utilize the symmetries of the matrix to reduce the time of multiplying a vector by this matrix." The columns (and rows) or this matrix are mutually orthogonal (complex-conjugate orthogonal). Observe that the elements of row 1 (not row 0) of the Fourier Matrix are points that uniformly distributed around the unit circle in the complex plane and complete one revolution of the unit circle. On the other hand, the elements of row 2 are equally distributed in the unit circle but complete 2 revolutions of the circle. The elements in row 4 complete 4 revolutions of the unit circle. Finally, one can argue that the elements of row 0 perform 5 revolutions of the unit circle (a single point repeating itself five times).

Exam Questions:

1. For what kind of functions is a Fourier Series defined? Write down the formula for the Fourier Series expansion (the “real” expansion, not the “complex” expansion) of a function that can be expanded using Fourier Series.
2. What is an even function? What is an odd function? Show that a function can be expressed as the sum of its even and odd parts.
3. How can you find the even and odd parts of a periodical function using the Fourier Series expansion? Explain the logic behind your answer.
4. What is Gibb’s effect as applied to Fourier Series?
5. What is Runge’s phenomenon?
6. Explain how a function $f(t)$ defined in the interval $0 \leq t < T$ can be periodically extended as an even function. Then explain the advantages of doing this type of periodical extension into an even function. Finally, describe how will the Fourier Series equation will change for this type of extended function. Make sure to clearly explain the logic behind each of your statements.
7. Write down the equation for the Complex Fourier Series and, assuming an expansion of a real, periodic function f(t), explain how the coefficients of the Complex Fourier Series expansion are related to the coefficients of the real Fourier Series expansion. Showing the formulas is not good enough, you must explain all the mathematical steps necessary to arrive at your final relationships.
8. For a Complex Fourier Series expansion of a real, periodic function, clearly explain why the coefficients C_n and C_{-n} are complex conjugates of each other.
9. A common plot used in frequency analysis is the plot of $\|C_n\|^2$ vs. ω_n, also known as the power-spectrum plot. Explain why this plot must be symmetric about the y-axis.
10. Write down Parseval’s Theorem as applied to the Complex Fourier Series expansion. Then clearly explain why this theorem must be true.
11. Find the length of the vector $e^{jn\omega_o t}$ in the interval $0 \leq \mathrm{t} < T$, where n is an integer number.
12. Find the dot product between vectors $e^{jn\omega_o t}$ and $e^{jm\omega_o t}$ where both n and m are integers, with $n \neq m$. Find the exact numerical value of the results and clearly explain the logic behind your calculations.
13. A common plot used in frequency analysis is the plot of $\|C_n\|^2$ vs. ω_n, also known as the power-spectrum plot. Explain why this plot must be symmetric about the y-axis.
14. Starting from the Complex Fourier Series, show how to derive the Discrete Fourier Transform. Clearly explain the logic behind every step in your derivation.
15. Find the numerical value of the dot product $\left(e^{-j\omega_n k}, e^{-j\omega_n k}\right)$ where $e^{-j\omega_n k}$ represent a vector from the Discrete Fourier Transform, and $k = 0,1, \ldots (N-1)$. Clearly explain the logic behind each step in your solution process.
16. Show how given uniformly spaced data points sampled from a real function $f(t)$ can be interpolated using the DFT to obtain

$$f_{interp}(t) = \frac{1}{N}\left(C_0 + 2Re\left(\sum_{k=1}^{\frac{N-1}{2}}\left(C_k e^{jk\omega_o t}\right)\right)\right)$$

17. Prove that only $N/2$ coefficients of the DFT are needed to reconstitute the function.
18. Clearly explain the logic used to express the DFT as a matrix multiplication.

Nov. 9- Thursday
LECTURE 23: Discrete Fourier Series – continuation

1. Quick review of lecture 22.
2. Convolution in the time domain is the same as multiplication in the frequency domain. For N dimensional vectors h and f, one can write:

$$DFT(h * f)_k = DFT(h)_k DFT(f)_k$$

The proof of this is given in the note on “Convolution: Time shift and DFT Theorems - Proof” that shows how to use the DFT to perform the convolution process. This note uses the $*$ notation to denote the convolution process between the functions h and f. For instance,

$$x = h * f$$

This notation is also introduced in the note “Convolution: Continuous to Discrete”

3. Recall that

$$DFT(f)_n = \sum_{k=0}^{N-1} f_k e^{-j\omega_n k}$$

and

$$IDFT(C)_n = \frac{1}{N}\sum_{k=0}^{N-1} C_k e^{j\omega_n k}$$

However, in Mathcad, the commands icfft and cfft are defined as

$$icfft(f)_n \equiv \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} f_k e^{-j\omega_n k}$$

and

$$cfft(C)_n \equiv \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} C_k e^{j\omega_n k}$$

so it follows that

$$DFT(f)_n = \sqrt{N}\cdot icfft(f)_n$$

and

$$IDFT(C)_n = \frac{1}{\sqrt{N}}\cdot cfft(C)_n$$

so

$$DFT(h * f)_k = DFT(h)_k \cdot DFT(f)_k = N \cdot icfft(h)_k \cdot icfft(f)_k$$

or, using vectorization notation (crtl – or ctrl minus in Mathcad),

$$DFT(h * f) = \overrightarrow{DFT(h)\cdot DFT(f)} = N \cdot \overrightarrow{\imath cfft(h)\cdot \imath cfft(f)}$$

and if

$$x = h * f$$

then

$$x = IDFT(DFT(h * f)) = IDFT\left(\overrightarrow{DFT(h) \cdot DFT(f)}\right) = \sqrt{N} \cdot cfft\left(\overrightarrow{icfft(h) \cdot icfft(f)}\right)$$

Finally, observe that the above can also work by exchanging the $cfft$ and the $icfft$ to obtain

$$x = \sqrt{N} \cdot icfft\left(\overrightarrow{cfft(h) \cdot cfft(f)}\right)$$

4. The note on “Convolution - Mathcad cfft and icfft” shows how to use Mathcad commands to perform the convolution process using the Mathcad cfft and icfft commands.
5. Beware: the fft command in Mathcad only accepts vectors with number of points equal to a power of 2. However, the cfft command accepts vectors with any number of points.
6. The note on “Convolution - Multiplying Polynomials” demonstrates how to use the cfft & icfft command to multiply polynomials using the convolution process, following the same example used in class.
7. The note on “Convolution - Moving Average Filter example” demonstrates how to use the cfft & icfft command to perform a moving average filter using the convolution process, following a similar example to that used in class.

Exam Questions:

1. Clearly show how multiplying two polynomials can be expressed as a convolution process. Then explain how the Discrete Fourier Transform can used to perform the polynomial multiplication. You must be able to completely explain the logic behind this process.
2. Clearly show how a moving average filter can be implemented using on the convolution. Then explain how the Discrete Fourier Transform can used to perform this task. You must be able to completely explain the logic behind this process.

Nov. 14- Tuesday
LECTURE 24: Eigenvalues & Eigenvectors

Summary:

1. Read Chapter 6 in your textbook.
2. The “leakage” effect was explained as a result of the convolution process in the frequency domain due to the multiplication of a boxcar (truncation) function and the function being analyzed.
3. Since the beginning of the course we have been studying how a coordinate system can used to simplify the solution process of a system of equations. The Gram-Schmidt orthogonalization process was used to find a set of orthogonal vectors that were used as coordinates. The Legendre, Chebyshev, and Sinusoidal (Fourier Series) functions were examples of coordinates in continuous systems that could be used to simplify/analyze a system of equations. The question arises on whether there is a coordinate system that is particularly well suited for a given system of equations. This question is partially answered by the matrix eigenvalue problem.
4. "Eigen" means "special" or "unique" or "singular".

5. An eigenvector x of a matrix A is a special vector that do not change direction when multiplied by A: $Ax = \lambda x$. The scalar λ can be considered a "stretching" factor.
6. The magnitude of the eigenvectors is not unique. If $Ax = \lambda x$ then $A(cx) = \lambda(cx)$ where c is a scalar. Thus (cx) is also and eigenvector.
7. For a *square* matrix A, the problem $Ax = \lambda x$ is called the matrix eigenvalue problem.
8. If a nxn matrix A has n independent eigenvectors then it is possible to transform the matrix into a diagonal matrix as follows. First write

$$[Ax_1\ Ax_2 \ldots Ax_n] = [\lambda_1 x_1\ \lambda_2 x_2 \ldots \lambda_n x_n]$$

this equation can be expressed as

$$AS = S\Lambda$$

where

$$S = [x_1\ x_2 \ldots x_n]$$

and

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

It follows that

$$\Lambda = S^{-1}AS$$

Note that S^{-1} exists because the eigenvectors of A are independent. Also note that

$$A = S\Lambda S^{-1}$$

9. If a nxn matrix A has n independent eigenvectors then the problem $Ax = b$ can be studied through a change of coordinates based on the eigenvectors of A. Let the coordinate transformation be represented by $x = Sy$, where, as described previously, S is a matrix containing the eigenvectors of A. Then the problem

$$Ax = b$$

can be expressed as

$$ASy = b$$

Multiplying both sides of the equation by S^{-1} yields

$$S^{-1}ASy = S^{-1}b$$

or

$$\Lambda y = \bar{b} \quad \text{where} \quad \bar{b} = S^{-1}b$$

This means that

$$\begin{array}{c} \lambda_1 y_1 = \bar{b}_1 \\ \lambda_2 y_2 = \bar{b}_2 \\ \vdots \\ \lambda_n y_n = \bar{b}_n \end{array} \quad \text{or} \quad \begin{array}{c} y_1 = \bar{b}_1/\lambda_1 \\ y_2 = \bar{b}_2/\lambda_2 \\ \vdots \\ y_n = \bar{b}_n/\lambda_n \end{array}$$

10. Read the note on the course website on obtaining a set of first order differential equations from higher order differential equations.
11. Converting a higher order differential equation to a set of first order differential equations allows for the use of the matrix eigenvalue problem to obtain a coordinate system that may yield a better way to find the solution of a differential equation. For instance, the differential equation

$$\underbrace{\dddot{y}}_{\dot{y}_2} + a \cdot \underbrace{\ddot{y}}_{y_2} + b \cdot \underbrace{\dot{y}}_{y_1} + c \cdot \underbrace{y}_{y_0} = u(t)$$

can be expressed as

$$\frac{d}{dt}\underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -c & -b & -a \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{B} u(t)$$

Matrix A is known as the "**Companion matrix**" for the characteristic equation $(s^3 + as^2 + bs + c = 0)$ of the differential equation.

12. In matrix notation one can write

$$\dot{x} = Ax + Bu$$

If the matrix A is not diagonal, then the problem is said to be "coupled" in the sense that the derivatives of, say, $x_i(t)$ depend on other x's other than $x_i(t)$. In such cases it is said that $x_i(t)$ is coupled to other x's. If the matrix A have independent eigenvectors, then it is possible to simplify the problem by "diagonalizing" or "decoupling" the system of equation by transforming the problem such that the resulting transformed counterpart of A is a diagonal matrix. This can be achieved as follows. First the eigenvectors of A are assembled as columns in a matrix S. Then, the following coordinate transformation

$$x = Sy$$

is used to obtain

$$\frac{d}{dt}Sy = ASy + Bu$$

or

$$S\frac{d}{dt}y = ASy + Bu$$

premultiplying both sides of the equation by S^{-1} yields,

$$\frac{d}{dt}y = \underbrace{[S^{-1}AS]}_{\Lambda}\, y + \underbrace{[S^{-1}B]}_{\bar{B}}\, u$$

From the bullet 8 in this lecture note, it is known that

$$\Lambda \ = \ S^{-1}AS$$

Defining

$$\bar{B} = S^{-1}B$$

and substituting into the differential equation, one obtains

$$\frac{d}{dt}y = \Lambda y + \bar{B}u$$

Recalling that

$$\Lambda \ = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

the differential equation becomes

$$\frac{d}{dt}y_1 = \lambda_1 y_1 + [\bar{B}u]_1$$

$$\vdots$$

$$\frac{d}{dt}y_n = \lambda_n y_n + [\bar{B}u]_n$$

The differential equations above (in y) are said to be "uncoupled" this means that the solution of any of differential equations do not depend on the solution of the rest. For instance, the solution for y_1 does not depend on $y_2, y_3, \dots, y_n$.

Using convolution, the solution of the i-th first order differential equation is given by

$$y_i(t) = e^{\lambda_i t}y_i(0) + \int_0^t h_i(t-\tau)f_i(t)d\tau$$

where $h_i(t)$ is the impulse response for the i-th uncoupled differential equation, and $f_i(t)$ is the input function for the i-th differential equation, i.e.,

$$f_i(t) = [\bar{B}u(\tau)]_i$$

The i-th impulse response function, $h_i(t)$, is the solution the i-th uncoupled differential equation with an unit impulse input:

$$\frac{d}{dt}h_i = \lambda_i h_i + \underbrace{\delta(t)}_{unit\ impulse}$$

Using Laplace Transforms, $\mathcal{L}(\delta(t)) = 1$, the Laplace Transform of the differential equation above is

$$sh_i = \lambda_i h_i + 1$$

or

$$h_i = \frac{1}{s - \lambda_i}$$

Applying the Inverse Laplace transform, the impulse response is found to be

$$h_i(t) = e^{\lambda_i t}$$

Substituting the impulse response result and the definition for $f_i(t)$ back into the solution yields,

$$y_i(t) = e^{\lambda_i t} y_i(0) + \int_0^t e^{\lambda_i (t-\tau)} [\bar{B}u(\tau)]_i d\tau$$

The array of solutions can be expressed in matrix form as

$$\underbrace{\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}}_{y(t)} = \underbrace{\begin{bmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{bmatrix}}_{e^{\Lambda t}} \underbrace{\begin{bmatrix} y_1(0) \\ y_2(0) \\ \vdots \\ y_n(0) \end{bmatrix}}_{y(0)} + \int_0^t \underbrace{\begin{bmatrix} e^{\lambda_1 (t-\tau)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n (t-\tau)} \end{bmatrix}}_{e^{\Lambda (t-\tau)}} \bar{B}u(\tau)d\tau$$

or in shortcut notation,

$$y(t) = \ e^{\Lambda t}\, y(0) + \int_0^t e^{\Lambda (t-\tau)} \bar{B}u(\tau)d\tau$$

Observe that $e^{\Lambda t}$ is defined as

$$e^{\Lambda t} \triangleq \begin{bmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{bmatrix}$$

The final solution $x(t)$ is found by transforming the coordinates back:

$$y = S^{-1}x \quad and \quad \bar{B} = S^{-1}B$$

to obtain

$$S^{-1}x(t) = \ e^{\Lambda t}\, S^{-1}x(0) + \int_0^t e^{\Lambda (t-\tau)}\, S^{-1}Bu(\tau)d\tau$$

or, pre-multiplying both sides of the equation by S

$$x(t) = \ \underbrace{\lfloor S e^{\Lambda t}\, S^{-1} \rfloor}_{e^{At}} x(0) + \int_0^t \underbrace{\lfloor S e^{\Lambda (t-\tau)} S^{-1} \rfloor}_{e^{A(t-\tau)}} Bu(\tau)d\tau$$

or in shortcut notation

$$x(t) = \underbrace{[e^{At}x(0)]}_{\substack{free\ response\\ a.k.a\\ zero-input\ response}} + \underbrace{\left[\int_0^t e^{A(t-\tau)}Bu(\tau)d\tau\right]}_{\substack{forced\ response\\ a.k.a\\ zero-state\ response}}$$

where e^{At} is the “matrix exponential” defined as

$$e^{At} \triangleq \; Se^{\Lambda t}\,S^{-1} = S\begin{bmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n t}\end{bmatrix}S^{-1}$$

Observe that the solution of the system of differential equations can be divided into the sum of two intermediate solutions: the free response (a.k.a zero-input response) and the forced response (a.k.a. zero-state response.) The free response is the response due to initial conditions assuming the input u(t) is zero. In contrast, the forced response is the response obtained by assuming the initial conditions are zero.

Exam Questions:

1. Clearly explain the logic of how the “leakage” effect appears when performing a frequency analysis of a truncated function.
2. Explain the problem of “aliasing” when under-sampling a function. (read http://www.svi.nl/AliasingArtifacts).
3. Explain how a moving average filter works and how it can be implemented using a matrix multiplication resulting in a convolution process.
4. Derive the fact that a nxn matrix A with n independent eigenvectors can be expressed as $A \; = \; S\Lambda S^{-1}$ where S is a matrix containing the eigenvectors of A. Your derivation must follow clear and logical steps supported by fundamental concepts.
5. Given a nxn invertible matrix A with n independent eigenvectors, show how a change in coordinate systems can be used to decouple the problem $Ax = b$. Use the decoupled system to find the solution x to the original problem $Ax = b$. As usual, your solution process must follow clear and logical steps supported by fundamental concepts.
6. Explain how to express the differential equation $\dddot{z} + a \cdot \ddot{z} + b \cdot \dot{z} + c \cdot z = u(t)$ as a matrix problem of the form $\dot{x} = Ax + Bu$.
7. Show how to use the matrix eigenvalue problem along with a coordinate transformation to decouple the equation $\dot{x} = Ax + Bu$. Your solution process must follow clear and logical steps supported by fundamental concepts.
8. Using the convolution integral, show how to find the complete solution to the following first order differential equation $\frac{d}{dt}y = \lambda y + f(t)$ with initial condition $y(0)$. Make sure to explain how to obtain the impulse response use in the convolution integral. Your solution process must follow clear and logical steps supported by fundamental concepts.
9. Use the answers to questions 3, 4, and 5 to explain the logical steps involved in finding the complete solution to the differential equation $\dddot{z} + a \cdot \ddot{z} + b \cdot \dot{z} + c \cdot z = u(t)$ with known initial conditions. Your solution must be based on free and forced responses and must rely on the concept of the matrix exponential e^{At}. Since the matrix exponential solution is a time varying vector, clearly explain how to plot $z(t)$ vs. t using the solution from the matrix exponential formula.

Nov. 16- Thursday
LECTURE 25: Eigenvalues & Eigenvectors, continuation

Summary:

1. Make sure to read Chapter 6 in your textbook.
2. Exhaustive review of Lecture 24
3. Read the pdf file on the Power Method for finding eigenvectors and eigenvalues. You will find this file on the Notes section of the course website.
4. Not discussed in class. In Lecture 24, the matrix exponential was defined for a square matrix with independent eigenvectors. The matrix exponential for any square matrix can be defined as a generalization of the exponential function. The Taylor Series expansion of the exponential function is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

So the matrix exponential is also defined as

$$e^{At} = \sum_{n=0}^{\infty} \frac{(At)^n}{n!}$$

The exponential function is useful in solving differential equations because

$$\left[\frac{d}{dt}\right] e^{at} = ae^{at}$$

i.e., the exponential function e^{at} is an "eigenvector" or "eigenfunction" of the derivative operator $\left[\frac{d}{dt}\right]$. The eigenvalue is a. Similarly,

$$\left[\frac{d}{dt}\right] e^{At} = Ae^{At}$$

It follows that the free (or zero-input) response of the matrix differential equation

$$\dot{x} = Ax + Bu(t)$$

is solved by considering

$$\dot{x}_{free} = Ax_{free}$$

$$x_{free}(t) = e^{At}C$$

where C is a $n \times 1$ constant vector. To see that this is indeed the solution, just substitute the solution back into the differential equation. To obtain the free response, apply the initial conditions to obtain $C = x(0)$.
The forced response is obtained, by convolution,

$$x_{forced}(t) = \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

and the complete, final solution is

$$x(t) = x_{free}(t) + x_{forced}(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

5. Not discussed in class. The roots of a polynomial can be obtained from the eigenvalues of the polynomial "Companion" matrix (to be defined below) as follows. Consider the solution (roots) of the polynomial equation

$$x^3 + ax^2 + bx + c = 0$$

This equation can be expressed in matrix form as

$$\underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -c & -b & -a \end{bmatrix}}_{companion\ matrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} = x \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}$$

Comparing to the matrix eigenvalue problem, $Az = \lambda z$ it is seen that matrix A corresponds to the companion matrix, the eigenvector z corresponds to the vector $\begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}$ and the eigenvalue λ corresponds to x. Since the value of x satisfying the polynomial equation corresponds to the value of x satisfying the corresponding matrix eigenvalue problem, it follows that the eigenvalues of the polynomial companion matrix correspond to the roots of the polynomial equation.

Let $\lambda_1, \lambda_2, \lambda_3$ correspond to the eigenvalues of the companion matrix. Then the eigenvectors can be assembled into a matrix S as follows

$$S = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ {\lambda_1}^2 & {\lambda_2}^2 & {\lambda_3}^2 \end{bmatrix}}_{Vandermonde\ Matrix}$$

Previously, it was shown that $S^{-1}AS = \Lambda$, so it follows that

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ {\lambda_1}^2 & {\lambda_2}^2 & {\lambda_3}^2 \end{bmatrix}}_{Vandermonde\ Matrix}^{-1} \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -c & -b & -a \end{bmatrix}}_{Companion\ Matrix} \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ {\lambda_1}^2 & {\lambda_2}^2 & {\lambda_3}^2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

So it is seen that the Vandermonde matrix diagonalizes the polynomial companion matrix. What is important here is that once the eigenvalues of the companion matrix are known, the eigenvectors can be immediately assembled without further calculations.

6. Polynomial Companion matrices also appear when expressing a higher order linear differential equation with constant coefficients as a system of first order differential equations. This was shown in Lecture 24.

7. Not discussed in class. Chapter 6 in your textbook shows how to find eigenvectors and eigenvalues by hand. The trick is to write $Ax = \lambda x$ as $Ax - \lambda \mathrm{I}x = 0$ and then factor x out to obtain
$(\mathrm{A} - \lambda\mathrm{I})x = 0$. Since the eigenvector x is not zero then it must lie in the nullspace of $(\mathrm{A} - \lambda\mathrm{I})$ implying that $(\mathrm{A} - \lambda\mathrm{I})$ is a singular matrix. The determinant of a singular matrix is zero so it follows that $\|\mathrm{A} - \lambda\mathrm{I}\| = 0$ (here the double vertical bar means determinant). Expanding the determinant $\|\mathrm{A} - \lambda\mathrm{I}\|$ will yield an n-th order polynomial in the variable λ. The equation $\|\mathrm{A} - \lambda\mathrm{I}\| = 0$ can be interpreted as solving for the roots of a polynomial. Once the roots (values of λ) are obtained, then the equation
$(\mathrm{A} - \lambda\mathrm{I})x = 0$ can be used for each root (λ) found to solve for the eigenvectors.

8. Not discussed in class. It is important to realize that the magnitude of the eigenvectors is not unique: if $Ax = \lambda x$ then $A(cx) = \lambda(cx)$ for any scalar c. So (cx) is also an eigenvector with the same direction but different magnitude. This implies that one needs to specify one more equation to find an eigenvector. For instance one could specify that the magnitude of the eigenvector be 1, i.e., $\|x\| = 1$. However, this is not the easiest choice to apply. A more convenient specification is to set one element of the eigenvector to 1. The other elements of the eigenvector can be obtained from the equation $(\mathrm{A} - \lambda\mathrm{I})x = 0$.

Example 1: $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

Step 1: evaluate $\|\mathrm{A} - \lambda I\|$: $\left\|\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}\right\| = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$

Step 2: solve for the roots: $\lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1) = 0$ then $\lambda_1 = 1$ and $\lambda_2 = 3$

Step 3: solve for the eigenvectors using the equation $(\mathrm{A} - \lambda I)x$:

assuming the first element of x is 1.

$$\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} 1 \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

From the first equation $a = \lambda - 2$

When $\lambda = \lambda_1 = 1$ then $a = 1 - 2 = -1$ and the eigenvector is $x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

When $\lambda = \lambda_2 = 3$ then $a = 3 - 2 = 1$ and the eigenvector is $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Example 2: $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$

Step 1: evaluate $\|\mathrm{A} - \lambda I\|$:

$$\left\|\begin{pmatrix} 2-\lambda & 1 & 0 \\ 1 & 3-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{pmatrix}\right\| = (2-\lambda)[(3-\lambda)(2-\lambda) - 1] - (2-\lambda) = -\lambda^3 + 7\lambda^2 - 14\lambda + 8$$

Step 2: solve for the roots:

$$-\lambda^3 + 7\lambda^2 - 14\lambda + 8 = -(\lambda - 4)(\lambda - 2)(\lambda - 1) = 0 \text{ then } \lambda_1 = 1,\ \lambda_2 = 2, \text{and } \lambda_3 = 4$$

Step 3: solve for the eigenvectors using the equation $(\mathrm{A} - \lambda I)x$:
assuming the first element of x is 1.

$$\begin{bmatrix} 2-\lambda & 1 & 0 \\ 1 & 3-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} 1 \\ a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From the first equation $a = -(2 - \lambda)$
From the second equation $b = -(3 - \lambda)(\mathrm{a}) - 1 = (3 - \lambda)(2 - \lambda) - 1$

When $\lambda = \lambda_1 = 1$ then $a = -1$, $b = 1$, and the eigenvector is $x_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

When $\lambda = \lambda_2 = 2$ then $a = 0$, $b = -1$, and the eigenvector is $x_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

When $\lambda = \lambda_3 = 4$ then $a = 2$, $b = 1$, and the eigenvector is $x_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

9. Consider two $n \times n$ real symmetric matrices A and B. The generalized eigenvalue problem of the pair (A, B) is

$$Ax = \lambda Bx$$

where x is the generalized eigenvector and λ is the generalized eigenvalue. Next it is shown that the the generalized eigenvectors are orthogonal "with respect to matrices A and B".

Suppose λ_i and λ_j are two generalized eigenvalues and assume $i \neq j$ and $\lambda_i \neq \lambda_j$, i.e., the generalized eigenvalues are **distinct**. It follows that,

$$Ax_i = \lambda_i Bx_i$$

and

$$Ax_j = \lambda_j Bx_j$$

Premultiplying the first equation by ${x_j}^T$ and the second equation by ${x_i}^T$ results in

$${x_j}^T Ax_i = \lambda_i {x_j}^T Bx_i \qquad (1)$$
$${x_i}^T Ax_j = \lambda_j {x_i}^T Bx_j \qquad (2)$$

Note that every 1x1 matrix is symmetric. Since ${x_i}^T Ax_j$ is 1x1, it follows that

$${x_i}^T Ax_j = \left({x_i}^T Ax_j\right)^T = {x_j}^T A^T x_i = {x_j}^T Ax_i$$

similarly,

$$x_i{}^T B x_j = x_j{}^T B x_i$$

Using the above symmetry conditions equations (1) and (2) can be expressed as,

$$\begin{aligned} x_j{}^T A x_i &= \lambda_i x_j{}^T B x_i \\ x_j{}^T A x_i &= \lambda_j x_j{}^T B x_i \end{aligned}$$

Subtracting the bottom equation from the top one results in,

$$0 = (\lambda_i - \lambda_j) x_j{}^T B x_i$$

Since $\lambda_i \neq \lambda_j$ it follows that

$$x_j{}^T B x_i = 0$$

and using the above result with equation (1) it follows that

$$x_j{}^T A x_i = 0$$

Thus, it is seen that x_i and x_j are orthogonal with respect to matrices A and B.

For the case where $i \neq j$ but $\lambda_i = \lambda_j$, i.e., repeated eigenvalues, one can slightly modify one of the diagonal element elements of matrices A or B by adding a small number ε so as to try to separate the eigenvalues. Since the matrices are still symmetric, the eigenvectors will be orthogonal. On the limit as $\varepsilon \to 0$ the eigenvalues will approach the same value, but the eigenvectors will remain orthogonal. A more complete proof can be found using "Schur's lemma" which states that <u>any</u> square matrix A can be converted into an upper triangular matrix using an orthogonal matrix U, the upper triangular matrix is given by $U^T A U$. If A is further assumed to be real symmetric then $(U^T A U)^T = U^T A U$, i.e., the upper triangular matrix $U^T A U$ is symmetric. But the only way an upper triangular matrix can be symmetric is for it to be a diagonal matrix. It follows that matrix U diagonalizes the real symmetric matrix A, and the eigenvectors of A are the columns of U. Schur's lemma is easy to prove. Time permitting this will be explained in class at a later time.

1. Assembling all the generalized eigenvectors as columns of a matrix S,

$$S = [x_1 \;\; x_2 \;\; \cdots \;\; x_n]$$

provides a way to diagonalize both matrices A and B as follows,

$$S^T A S = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} A [x_1 \;\; x_2 \;\; \cdots \;\; x_n] = \begin{bmatrix} x_1^T A x_1 & \underbrace{x_2^T A x_1}_{0} & \cdots & \underbrace{x_n^T A x_1}_{0} \\ \underbrace{x_1^T A x_2}_{0} & x_2^T A x_1 & \cdots & \underbrace{x_n^T A x_2}_{0} \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{x_1^T A x_n}_{0} & \underbrace{x_2^T A x_n}_{0} & \cdots & x_n^T A x_n \end{bmatrix} = \begin{bmatrix} x_1^T A x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n^T A x_n \end{bmatrix}$$

Similarly,

$$S^T B S = \begin{bmatrix} x_1^T B x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n^T B x_n \end{bmatrix}$$

2. Observe that if $B = I$, then, as discussed before, the generalized eigenvalue problem $Ax = \lambda Bx$ reduces to the matrix eigenvalue problem $Ax = \lambda x$. In particular, assuming the eigenvectors of A are normalized, then $S^T BS$ becomes $S^T IS = S^T S = I$. This means that the eigenvectors are orthogonal and S^T is the inverse of S, i.e., $S^{-1} = S^T$. In this case S is said to be an orthogonal matrix. This proves that the eigenvectors of a real symmetric matrix (A) are orthogonal.
3. An important application of the generalized eigenvalue problem is modal analysis. In that case one considers the problem

$$M\ddot{x} + Kx = f(t)$$

where M and K are real symmetric nxn matrices: the mass matrix and the stiffness matrix. By the generalized eigenvalue problem, i.e., let $A = M$ and $B = K$, then

$$S^T M S = \begin{bmatrix} x_1^T M x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n^T M x_n \end{bmatrix} \text{ and } S^T K S = \begin{bmatrix} x_1^T K x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n^T K x_n \end{bmatrix}$$

The diagonal terms $m_i = x_i^T M x_i$ and $k_i = x_i^T K x_i$ for $i = 1,2,\cdots,n$ are known as the modal masses and modal stiffnesses.
The differential equation can be solved by the coordinate transformation $x = Sy$ as follows. First substitute $x = Sy$.

$$MS\ddot{y} + KSy = f(t)$$

pre-multiplying by S^T

$$S^T MS\ddot{y} + S^T KSy = S^T f(t)$$

Let $\bar{f}(t) = S^T f(t)$. Then, the above equation is decoupled and can be written as

$$\begin{gathered} m_1\ddot{y}_1 + k_1 y_1 = \bar{f}_1(t) \\ m_2\ddot{y}_2 + k_2 y_2 = \bar{f}_2(t) \\ \vdots \\ m_n\ddot{y}_n + k_n y_n = \bar{f}_n(t) \end{gathered}$$

Each equation can be solved separately using the free and forced responses and the initial conditions $y(0) = S^T x(0)$. For instance, if the input excitation is zero, i.e., $\bar{f}_i(t) = 0$ and the initial velocity is zero $\dot{y}_i(0) = 0$, then $y_i(t) = y_i(0)\cos(\omega_{n_i} t)$ where $\omega_{n_i} = \sqrt{k_i/m_i}$.

The final solution is given by transforming back to x coordinates:

$$x(t) = Sy(t) = S^{\langle 1 \rangle} y_1(t) + S^{\langle 2 \rangle} y_2(t) + \cdots S^{\langle n \rangle} y_n(t)$$

This means that the vector $x(t)$ is a linear combination of the columns of S. The columns of S are called "mode shapes" and represent different "modes" of vibrations: the final solution is a linear combination of n modes of vibration. Each mode of vibration represents a specified motion pattern for the masses in the system.

Exam Questions:

1. Explain how Taylor Series can be used to define the Matrix Exponential, then show how the definition of the Matrix Exponential can be used to solve a system of first order differential equations of the form $\dot{x} = Ax + Bu(t)$.
2. Explain how the problem of finding the roots of a polynomial of the form $x^3 + ax^2 + bx + c = 0$ can be converted into a matrix eigenvalue problem. Make sure to thoroughly provide the logic behind every step of your explanation.
3. Using clear and logical steps supported by fundamental concepts, explain how to find the eigenvalues and eigenvectors of a small 2x2 or 3x3 matrix. Be prepared to find the eigenvalues and eigenvectors for a given 2x2 or 3x3 matrix.
4. Using clear and logical reasoning, supported by fundamental concepts, explain how the power method to find eigenvectors and eigenvalues work.
5. Given two $n \times n$ real symmetric matrices A and B, prove that the generalized eigenvectors for this pair of matrices are orthogonal. Then use this fact to provide a transformation that can diagonalize both matrices A and B simultaneously. As usual, your solution process must follow clear and logical steps supported by fundamental concepts.
6. Using clear and logical steps supported by fundamental concepts, explain how to use the generalized matrix eigenvalue problem to solve the matrix differential equation $M\ddot{x} + Kx = 0$ where matrices M and K are real symmetric nxn matrices.

Nov. 21 - Tuesday
LECTURE 26: Principal Component Analysis (PCA)

1. Principal Component Analysis (PCA) can be construed as the process of finding eigenvectors for a data matrix. The main idea in PCA is to find a subspace (think plane or "hyperplane") that best aligns with the data in a data matrix. This is done by creating a real symmetric matrix using the data matrix as follows. Suppose that the data is obtained from an experiment where m distinct variables are collected at n intervals of time. This will produce an nxm matrix D of measurements (data), each column representing the measurements of a one of the variables. For instance, the columns of D can represent (1) pressure, (2) temperature, (3) density, (4) internal energy, and (5) entropy: imagine the thermo table. In this example the data is five dimensional. Every line (row) in the table is a new point. Together, all the points form a "cloud" or cluster of data points. The objective is to find the orientation (alignment) of the cloud of data points. The analysis is as follows.

   First translate the center of the cluster of data points to the origin by calculating the mean (average) for each column of D and then subtracting the means from the respective columns and label the new matrix M. Then consider a mx1 unit vector x. This vector will be aligned with the data if the direction of x is chosen such that the magnitude of $z = Mx$ is as large as possible. Observe that z is a nx1 vector with the i-th element being equal to the dot product between the i-th row of M and vector x. Since $z^T z$ is the square of the magnitude of z maximizing the magnitude of z is the same as maximizing $z^T z$.

$$z^T z = (Mx)^T(Mx) = x^T(M^T M)x$$

Next let $A = M^T M$. Since A is real symmetric, the eigenvectors of A are orthogonal and A may be expressed as $A = S\Lambda S^T$ where S is a nxn matrix with columns corresponding to orthonormal eigenvectors of A.

$$A = S\Lambda S^T = [S^{\langle 1\rangle} \quad \cdots \quad S^{\langle n\rangle}]\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \lambda_n\end{bmatrix}\begin{bmatrix}S^{\langle 1\rangle^T}\\ \vdots \\ S^{\langle n\rangle^T}\end{bmatrix}$$

or

$$A = S\Lambda S^T = \lambda_1 S^{\langle 1\rangle} S^{\langle 1\rangle^T} + \cdots + \lambda_n S^{\langle n\rangle} S^{\langle n\rangle^T}$$

Suppose the first two eigenvalues are several orders of magnitude larger than the rest. Then the data will be predominantly aligned with the plane (subspace) spanned by the eigenvectors corresponding to these two eigenvalues. In that case we may just project the data onto that plane (subspace) as follows. First observe that $S^{\langle 0\rangle}$ and $S^{\langle 1\rangle}$ can be thought of as the $\hat{\imath}$ and $\hat{\jmath}$ unit vectors:

$$\hat{\imath} = S^{\langle 0\rangle} \quad \text{and} \quad \hat{\jmath} = S^{\langle 1\rangle}$$

The projected data matrix is then

$$\underbrace{M_{filtered}}_{n\times m} = \underbrace{(M\hat{\imath})}_{n\times 1}\,\underbrace{(\hat{\imath}^T)}_{1\times m} + \underbrace{(M\hat{\jmath})}_{n\times 1}\,\underbrace{(\hat{\jmath}^T)}_{1\times m} = M[\hat{\imath} \quad \hat{\jmath}]\begin{bmatrix}\hat{\imath}^T\\ \hat{\jmath}^T\end{bmatrix} = MS_r{S_r}^T \quad \text{where} \quad S_r = [S^{\langle 0\rangle} \quad S^{\langle 1\rangle}]$$

where $M\hat{\imath}$ represents the dot product between each row of M, i.e., each data point, and the unit vector $\hat{\imath}$. The transpose is used to convert the results from column into row form as the data is given in row form (recall that each data point is a row of M).

Plotting the vectors $(M\hat{\jmath})$ vs. $(M\hat{\imath})$ will produce a picture of the data projected onto the principal plane. Any deviation of the data from the principal plane (subspace) will be small because the first two eigenvalues are orders of magnitude larger than the other eigenvalues. In many situations these deviations from the principal subspace are considered noise: the difference between M and $M_{filtered}$ should be small (this can be used as a verification!). Matrix $M_{filtered}$ can be thought of as a “filtered” or “corrected” version of matrix M. The equation for $M_{filtered}$ can also be expressed as

$$M_{filtered} = MS_r{S_r}^T \quad \text{where} \quad S_r = [S^{\langle 0\rangle} \quad S^{\langle 1\rangle}]$$

2. The process for Principal Component Analysis (PCA) can be summarized as follows

   (1) First construct a nxm matrix D of measurement data (each row of D corresponds to a m-dimensional data point)
   (2) Then strip the mean out of each column of D to construct a matrix M.
   (3) It is often desirable to normalize each column of M, specially if the columns of M represent measurements with different units or very different magnitudes.
   (4) Find the eigenvalues and associated eigenvectors of the real symmetric matrix $M^T M$.
   (5) Select the eigenvectors associated with the largest eigenvalues.
   (6) Project the data onto each of the eigenvectors selected in step (5). For instance, suppose that
   $S^{\langle 0\rangle}$ and $S^{\langle 1\rangle}$ are two eigenvectors associated with the largest eigenvalues. Then set $x = MS^{\langle 0\rangle}$ and $y = MS^{\langle 1\rangle}$. Plotting x vs. y will provide a plot of the “principal components” of

the data matrix M. Of course, in other examples one may consider keeping more than two components, i.e., eigenvectors.

(7) Matrix M can be filtered as follows: $M_{filtered} = MS_rS_r{}^T$ where $S_r = [S^{\langle 0\rangle} \quad S^{\langle 1\rangle}]$. That is $M_{filtered}$ represents the majority of information contained in matrix M, but it excludes the relatively small changes (often due to random measurement noise) in the directions perpendicular to $S^{\langle 0\rangle}$ and $S^{\langle 1\rangle}$.

3. The directions (eigenvectors) associated with the negligible (orders of magnitude smaller) eigenvalues are not useless. They can be used to obtain equations relating some of the variables with respect to the remaining variables (think of the columns of M as variables). Following the previous example, let $S^{\langle 2\rangle}$, $S^{\langle 3\rangle}$, and $S^{\langle 4\rangle}$ represent eigenvectors associated with negligible eigenvalues. Then define

$$\underbrace{F}_{5\times 3} = \begin{bmatrix} \underbrace{S^{\langle 2\rangle}}_{5\times 1} & \underbrace{S^{\langle 3\rangle}}_{5\times 1} & \underbrace{S^{\langle 4\rangle}}_{5\times 1} \end{bmatrix}$$

Let $[a \ \ b \ \ c \ \ d \ \ e]$ represent variables identifying each of the columns of the M matrix. Since $MF \cong 0$, it follows that the variables associated with each of the columns of M, say, $[a \ \ b \ \ c \ \ d \ \ e]$, are approximately perpendicular to the columns of F, i.e.,

$$\underbrace{[a \ \ b \ \ c \ \ d \ \ e]}_{1\times 5} \underbrace{F}_{5\times 3} \cong \underbrace{[0 \ \ 0 \ \ 0]}_{1\times 3}$$

Transposing both sides of the equation,

$$F^T \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

There are three equations (three rows), implying that three of the variables can be solved in terms of the remaining two. For instance a, b, and c can be expressed as functions of d and e as follows. First split matrix F^T in two as follows

$$\underbrace{F^T}_{3\times 5} = \begin{bmatrix} \underbrace{F_1}_{3\times 3} & \underbrace{F_2}_{3\times 2} \end{bmatrix}$$

Then the previous equation can be expressed as

$$\begin{bmatrix} \underbrace{F_1}_{3\times 3} & \underbrace{F_2}_{3\times 2} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\underbrace{F_1}_{3\times 3} \underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{3\times 1} + \underbrace{F_2}_{3\times 2} \underbrace{\begin{bmatrix} d \\ e \end{bmatrix}}_{2\times 1} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{3\times1} = -\overbrace{\underbrace{{F_1}^{-1}}_{3\times3}\ \underbrace{F_2}_{3\times2}}^{3\times2}\underbrace{\begin{bmatrix} d \\ e \end{bmatrix}}_{2\times1}$$

The above represents three equations, each equation with input variables d and e. For instance, if

$$-{F_1}^{-1}F_2 = \begin{bmatrix} \alpha_0 & \beta_0 \\ \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$$

the three equations will be

$$a = \alpha_0 d + \beta_0 e$$
$$b = \alpha_1 d + \beta_1 e$$
$$c = \alpha_2 d + \beta_2 e$$

To verify the above equations, one can use the data from matrix $M_{filtered}$. For this example, one can try to use the data from any two columns (corresponding to the selected "independent" variables) to recreate the data from the remaining columns. This will only be possible if the resulting matrix F_1 is invertible. If not, another choice of "independent" variable (columns) must be made.

Recall that the "cloud" of data was originally shifted to the origin by subtracting the average from each column. The above equations apply to the shifted data. In order to apply the equations back to the original data one needs to undo the shift as follows. First define

$$a = \bar{a} - a_{avg}$$
$$b = \bar{b} - b_{avg}$$
$$c = \bar{c} - c_{avg}$$
$$d = \bar{d} - d_{avg}$$
$$e = \bar{e} - e_{avg}$$

where a_{avg}, b_{avg}, c_{avg}, d_{avg}, and e_{avg} are the averages of the columns of D. Substituting back into the previous three equations one obtains

$$\bar{a} = \alpha_0\bar{d} + \beta_0\bar{e} + (a_{avg} - \alpha_0 d_{avg} - \beta_0 e_{avg})$$
$$\bar{b} = \alpha_1\bar{d} + \beta_1\bar{e} + (b_{avg} - \alpha_1 d_{avg} - \beta_1 e_{avg})$$
$$\bar{c} = \alpha_2\bar{d} + \beta_2\bar{e} + (c_{avg} - \alpha_2 d_{avg} - \beta_2 e_{avg})$$

Again add the corresponding averages back to the columns of matrix $M_{filtered}$ to create a new "filtered" version of matrix D and use the resulting filtered data to check the equations above.

The steps used are then:

1) Associate a variable name with each column of matrix D. In this problem the variable names are a,b,c, d, and e but in other applications these variable names depend on the data acquired. For example, the variable names may be stress, strain, temperature, density, etc.

2) The same variable names (and in the same order) are used for the columns of matrix M, but recall that, in this case, these variables are stripped of their mean values and may be normalized.
3) Based on the magnitudes of the eigenvalues of $M^T M$ decide how many principal directions (eigenvectors associated with the larger eigenvalues) should be used. Use the rest of the eigenvectors (not the principal directions) as columns of a matrix labeled F.
4) Observe from the previous example that the columns of F^T correspond to the same variable names (and in the same order) as the columns of matrix M. Make sure to keep track of the variable names associated with the columns of F^T.
5) Let p be the number of principal directions and n be the total number of variables, then it ***may*** be possible to obtain approximate expressions for $(n-p)$ of the variables as functions of the remaining p variables. First, choose the p variables that will be used as the independent variables. Then extract the columns of matrix F^T associated with the choice of independent variables and assemble these columns in a matrix labeled F_2. The remaining columns of F^T are then assembled as columns of a matrix labeled F_1. Again, it is important to keep track of the variable name associated with of each column of F_1 and F_2.
6) If matrix F_1 is invertible then the choice of independent variables will work. Otherwise, a different choice of independent variables must be made so as to ensure that F_1 is invertible.
7) The equations relating the dependent variables to the independent variables is $(vector\ of\ dependent$ variables$) = -{F_1}^{-1}F_2(vector\ of\ independent\ variables)$
8) Modify the results of step 7 to account for the fact that the variable names correspond to variables that have been stripped of their mean values and possibly normalized.

Exam Questions:

1. The main idea behind Principal Component Analysis is to find the subspace that best aligns with the data. Assuming that each row of a nxm matrix M is a m dimensional data point M, clearly explain how to find the best subspace that aligns with the "cloud" of data in this matrix. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.
2. Assuming that each row of a nxm matrix M is a m dimensional data point, using eigenvalue/eigenvector analysis, show each logical step (explain why and how for each step) needed to clean up the noise in the data in this matrix to produce a new matrix, $M_{filtered}$.
3. Given a data matrix M, where each column of the matrix represents a variable and each row represents a data point. Explain how to use the ideas behind Principal Component Analysis to find an equation relating some of the variables (user selected dependent variables) as functions of the rest of the variables (user selected independent variables.) Also, clearly explain the limitations in the choices of dependent and independent variables. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.

Nov. 28 - Tuesday
LECTURE 27: Final Lecture. Singular Value Decomposition (SVD)

1. The Final Exam will be a take home exam. An email will be sent to the class when the Final Exam is posted online later today. Please turn in your worked-out final exam to me (at my office for on-campus students) by 11 am on Thursday, Dec. 7.
2. Recall that the eigenvalue decomposition can be used only on square matrices and that real symmetric matrices are useful because they produce orthogonal eigenvectors which can be conveniently used in coordinate transformations. A less restrictive decomposition which yields two set of orthogonal matrices (i.e., columns are normalized and mutually orthogonal) and a diagonal matrix and applies to rectangular as well as square matrices is the Singular Value Decomposition (SVD):

$$\underbrace{A}_{n\times m} = \underbrace{U}_{(n\times n)}\ \underbrace{S}_{(n\times m)}\ \underbrace{V^T}_{(m\times m)}$$

where U and V are $n \times n$ and $m \times m$ orthogonal matrices and S is a $n \times m$ diagonal matrix, i.e., the elements not on the main diagonal of S are zero. The elements on the main diagonal of S are positive or zero and are called singular values. The SVD expansion can be seen in more detail below

$$A = \begin{bmatrix} U^{\langle 1\rangle} & U^{\langle 2\rangle} & \cdots & U^{\langle n\rangle}\end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma_m \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} V^{\langle 1\rangle^T} \\ V^{\langle 2\rangle^T} \\ \vdots \\ V^{\langle m\rangle^T} \end{bmatrix} \quad \text{assuming } n > m$$

3. The U, S, and V matrices used in the singular values decomposition of a $n \times m$ matrix A can be found using the eigenvalue decomposition as follows. Consider the case when $n \geq m$ and compute

$$\underbrace{A^T}_{m\times n}\ \underbrace{A}_{n\times m} = \left(\underbrace{U}_{(n\times n)}\ \underbrace{S}_{(n\times m)}\ \underbrace{V^T}_{(m\times m)}\right)^T \left(\underbrace{U}_{(n\times n)}\ \underbrace{S}_{(n\times m)}\ \underbrace{V^T}_{(m\times m)}\right) = VS^T\,\underbrace{U^TU}_{I}\,SV^T = \underbrace{V}_{(m\times m)}\ \underbrace{S^TS}_{(m\times m)}\ \underbrace{V^T}_{(m\times m)}$$

Since A^TA is real symmetric, it must have orthogonal eigenvectors. The eigenvalue decomposition of A^TA is then

$$A^TA = E\Lambda E^T$$

where the columns of E are the m orthogonal eigenvectors of A^TA and Λ is a $m \times m$ diagonal matrix containing the corresponding eigenvalues. Comparing both equations,

$$VS^TSV^T = E\Lambda E^T$$

So it is seen that $V = E$ and $S^TS = \Lambda$. Since the diagonal terms of S^TS are ${\sigma_1}^2, {\sigma_2}^2, \cdots, {\sigma_m}^2$ and the diagonal terms of Λ are the eigenvalues of A^TA, it follows that

$$\sigma_k = \sqrt{\lambda_k} \quad \text{for} \quad k = 1,2,\dots,m$$

To find U observe that

$$\underbrace{A}_{n\times m}\underbrace{V}_{m\times m} = US\underbrace{V^TV}_{I} = \underbrace{U}_{n\times n}\underbrace{S}_{n\times m}$$

or

$$[AV^{\langle 1\rangle} \quad AV^{\langle 2\rangle} \quad \cdots \quad AV^{\langle m\rangle}] = [\sigma_1 U^{\langle 1\rangle} \quad \sigma_2 U^{\langle 2\rangle} \quad \cdots \quad \sigma_m U^{\langle m\rangle}]$$

The above equation implies that

$$U^{\langle k\rangle} = \frac{1}{\sigma_k} AV^{\langle k\rangle} \quad \text{for} \quad k = 1,2,\dots,r$$

where r is the rank of A and the singular values are assumed to be assembled in order of decreasing magnitude. The remaining columns $U^{\langle r+1\rangle}, U^{\langle r+2\rangle}, \cdots, U^{\langle n\rangle}$ (which span the left-nullspace of A) can be found using the Gram-Schmidt orthogonalization method applied to the set

$$\{U^{\langle 1\rangle}, U^{\langle 2\rangle}, \cdots, U^{\langle r\rangle}, e^{\langle 1\rangle}, e^{\langle 2\rangle}, \cdots, e^{\langle n\rangle}\}$$

where $e^{\langle 1\rangle}, e^{\langle 2\rangle}, \cdots, e^{\langle n\rangle}$ are the columns of the $n \times n$ identity matrix. Since $U^{\langle 1\rangle}, U^{\langle 2\rangle}, \cdots, U^{\langle r\rangle}$ are already orthogonal the Gram-Schmidt procedure starts by determining $U^{\langle r+1\rangle}$ from the part of $e^{\langle 1\rangle}$ that is orthogonal to $U^{\langle 1\rangle}, U^{\langle 2\rangle}, \cdots, U^{\langle r\rangle}$. If this does not work out, the process continues with $e^{\langle 2\rangle}$, etc.. until $U^{\langle r+1\rangle}$ is found. The remaining e vectors are used in a similar process to find the rest of the columns of U.

Finally, if $n < m$, repeat the methodology described above for A^T, i.e., find the SVD of A^T instead of A. When done, transpose the results to obtain the SVD of A. More explicitly, if $A^T = \bar{U}\bar{S}\bar{V}^T$ and $A = USV^T$, then $U = \bar{V}$, $S = \bar{S}^T$, and $V = \bar{U}$.

To see in more detail how the SVD works consider the expanded form

$$A = [U^{\langle 1\rangle} \quad U^{\langle 2\rangle} \quad \cdots \quad U^{\langle n\rangle}] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma_m \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} {V^{\langle 1\rangle}}^T \\ {V^{\langle 2\rangle}}^T \\ \vdots \\ {V^{\langle m\rangle}}^T \end{bmatrix} \quad \text{assuming } n > m$$

If, as an example, assume that $\sigma_3 = \sigma_4 = \cdots = \sigma_m = 0$ then

$$Ax = \sigma_1 U^{\langle 1\rangle}\left({V^{\langle 1\rangle}}^T x\right) + \sigma_2 U^{\langle 2\rangle}\left({V^{\langle 2\rangle}}^T x\right) + 0 \cdot U^{\langle 3\rangle}\left({V^{\langle 3\rangle}}^T x\right) + \cdots + 0 \cdot U^{\langle n\rangle}\left({V^{\langle m\rangle}}^T x\right)$$

Recall that x is part of the domain of A which corresponds to the rowspace and the nullspace. Focusing on the rowspace of A, it is seen that Ax cannot be zero as long as x has components along $V^{\langle 1\rangle}$ and/or $V^{\langle 2\rangle}$, which is equivalent as saying that $V^{\langle 1\rangle}$ and $V^{\langle 2\rangle}$ span the rowspace of A.

On the other hand, if x only has components along $V^{\langle 3\rangle}$ through $V^{\langle m\rangle}$ it is seen that $Ax = 0$, implying that vectors $V^{\langle 3\rangle}$ through $V^{\langle m\rangle}$ span the nullspace of A.
Looking at the co-domain of A, it is seen that since Ax is a linear combination of vectors $U^{\langle 1\rangle}$ and $U^{\langle 2\rangle}$ (recall that $\sigma_3 = \sigma_4 = \cdots = \sigma_m = 0$), these two vectors span the columnspace of A. Since $U^{\langle 3\rangle}$ through $U^{\langle n\rangle}$ are orthogonal to $U^{\langle 1\rangle}$ and $U^{\langle 2\rangle}$ and complete the space, they must span the left-nullspace of A.
Observe again, that $V^{\langle 1\rangle}$ and $V^{\langle 2\rangle}$ as well as $U^{\langle 1\rangle}$ and $U^{\langle 2\rangle}$ are associated with the non-zero singular values.

4. In general, the columns of V associated with the non-zero singular values form an orthonormal basis for the rowspace of A, and the columns of U associated with the non-zero singular values form an orthonormal basis for the columnspace of A
5. The rank of matrix A, i.e., the number of independent rows which is the same as the number of independent columns, is always equal to the number of nonzero singular values.
6. The importance of matrices with orthonormal columns is that projection matrices onto their column spaces are very simple. For instance, if matrix Q (not necessarily square) has orthonormal columns, then the projection into the columnspace of Q is just $P_{\mathcal{C}} = QQ^T$.
7. Projection matrices onto the four spaces of matrix A are readily found using the singular value decomposition of A. The columns of the V matrix associated with non-zero singular values can be used as the columns of matrix Q and the projection onto the rows space is then $P_{RS} = QQ^T$. If the columns of the V matrix associated with the zero singular values are used as the columns of matrix Q, then the projection onto the nullspace is then $P_{NS} = QQ^T$. Similarly, if the columns of the U matrix associated with non-zero singular values are used as the columns of matrix Q, then the projection onto the columnspace is then $P_{CS} = QQ^T$. Finally, if the columns of the U matrix associated with the zero singular values are used as the columns of matrix Q, then the projection onto the left-nullspace is then $P_{LN} = QQ^T$.
8. Since the SVD provides an orthonormal basis for the rowspace as well as for the columnspace of any matrix, it is an ideal tool to find the shortest least squares solution of the problem $Ax \neq b$. For instance, for the previous example where $\sigma_3 = \sigma_4 = \cdots = \sigma_m = 0$, removing the zero singular values from the SVD expansion as well as the vectors associated with the nullspace ($V^{\langle 3\rangle}$ through $V^{\langle m\rangle}$) and the leftnullspace ($U^{\langle 3\rangle}$ through $U^{\langle n\rangle}$) results in

$$\bar{A} = \underbrace{\begin{bmatrix} U^{\langle 1\rangle} & U^{\langle 2\rangle} \end{bmatrix}}_{n\times 2} \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} {V^{\langle 1\rangle}}^T \\ {V^{\langle 2\rangle}}^T \end{bmatrix}}_{2\times m} = \bar{U}\bar{S}\bar{V}^T$$

it follows that $Ax = \bar{A}x$ so the (inconsistent) matrix problem $Ax \neq b$ can be expressed as $\bar{A}x \neq b$ or

$$\bar{U}\bar{S}\bar{V}^T x \neq b$$

premultiplying by $\bar{U}^T$provides the least squares compromise (the inequality becomes an equality by finding components along the columnspace)

$$\bar{U}^T\bar{U}\bar{S}\bar{V}^T x = \bar{U}^T b$$

Since the columns of $\bar{U}$ are orthonormal then $\bar{U}^T\bar{U} = I$ and the above equation reduces to

$$\bar{S}\bar{V}^T x = \bar{U}^T b$$

premultiplying by $\bar{S}^{-1}$ results in

$$\bar{V}^T x = \bar{S}^{-1}\bar{U}^T b$$

The shortest solution requires x to lie on the rowspace of A meaning that x can be expressed as a linear combination of the columns of $\bar{V}$, i.e., $x = \bar{V}z$. Substituting above,

$$\bar{V}^T\bar{V}z = \bar{S}^{-1}\bar{U}^T b$$

but $\bar{V}^T\bar{V} = I$ because the columns of $\bar{V}$ are orthonormal and the above equation becomes

$$z = \bar{S}^{-1}\bar{U}^T b$$

The shortest least squares solution is finally revealed as,

$$x = \bar{V}z = \bar{V}\bar{S}^{-1}\bar{U}^T b$$

9. The matrix $\bar{V}\bar{S}^{-1}\bar{U}^T$ is known as the Moore-Penrose pseudo-inverse of $A = USV^T$.
10. An example of using the SVD in image compression is given in: http://www.uwlax.edu/faculty/will/svd/compression/index.html
11. In many applications none of the singular values are zero but some singular values are much smaller than the rest. In such cases, it is common to assume the smaller singular values are zero and apply the pseudo inverse to find the solution.
12. If the singular values are in decreasing order $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq \sigma_{r+1} \geq \cdots \geq \sigma_p$ where p is the smaller value between m and n, and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_p = 0$, then matrices $\bar{V}$ and $\bar{U}$ are obtained by keeping only the first r columns of the respective matrices V and U, and matrix $\bar{S}$ is an $r \times r$ matrix with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ along the diagonal.

13. Not discussed in class. A correlation filter based on the SVD can be used to filter some of the noise out of a data stream. The main idea is to assume the data stream is follows the following moving average (MA) filter equation

$$\hat{d}_{k+1} \cong c_0 d_k + c_1 d_{k-1} + \cdots + c_N d_{k-N}$$

where $\hat{d}_{k+1}$ is the filtered estimate of d_{k+1} using a moving, weighted average of the actual data $d_k, d_{k-1}, \cdots, d_{k-N}$ with weights $c_0, c_1, \cdots, c_N$. The coefficients c show how well $d_k, d_{k-1}, \cdots, d_{k-N}$ correlate with d_{k+1}, hence the name correlation filter. The above equation can be expressed in matrix form as follows

$$\underbrace{\begin{bmatrix} d_0 & d_1 & d_2 & \cdots & d_N \\ d_1 & d_2 & d_3 & \cdots & d_{N+1} \\ d_2 & d_3 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ d_p & d_{p+1} & \cdots & \cdots & d_{N+p} \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_N \end{bmatrix}}_{c} \cong \begin{bmatrix} d_{N+1} \\ d_{N+2} \\ d_{N+3} \\ \vdots \\ d_{N+p+1} \end{bmatrix}$$

Observe that the first column and the last row of the Hankel (a matrix constant along the anti-diagonals) matrix D above makes up a stream of $N_T = (N + p + 1\ )$ consecutive data points. Also observe that $p + 1$ is the number of rows and $N + 1$ is the number of columns of matrix D which is the same as the number of filter coefficients. Given a data stream of N_T consecutive data points, then $p = (N_T - N - 1)$ so the number of rows is dependent on the filter size. The SVD can be used to find the vector of coefficients. The idea is the same as before, perform the SVD of the Hankel matrix D, then plot the singular values against their indices and, using the plot, decide on a cut-off index that separates the dominant (larger) singular values from the rest. Keeping only the dominant singular values, and the associated rows of the U and V matrix create reduced dimension matrices $\overline{U}$, $\bar{S}$, and $\bar{V}$ to obtain a filtered data matrix

$$D_{filtered} = \overline{U}\bar{S}\bar{V}^T$$

The first column together with the last row of $D_{filtered}$ makes up the filtered data. There is no need to find the coefficient vector c, unless one plans to implement an FIR filter. Also note that by using the first column together with the last row of $D_{filtered}$ all the data will be filtered, however when using the moving average equation enough data points need to be available for the first weighted average to be calculated.

Observe that

$$D_{i,j} = d_{i+j} \text{ for } i = 0,1,\cdots,p \text{ and } j = 0,1,\cdots,N$$

A file entitled “CorrelationFilter.xmcd” has been posted in the Notes section of the course website. This file provides a Mathcad file on how to implement the correlation filter.

Exam Questions:

1. For the singular value decomposition $\underbrace{A}_{n\times m} = \underbrace{U}_{(n\times n)} \underbrace{S}_{(n\times m)} \underbrace{V^T}_{(m\times m)}$, explain how to find the matrices U, S, and V, assuming that $n \geq m$. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.
2. In the singular value decomposition $A = USV^T$, explain why matrices U and V are orthogonal matrices.
3. Given the singular value decomposition of matrix A, i.e., $A = USV^T$, explain how to find projection matrices into each of the four spaces of matrix A. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.
4. Show how to use the singular value decomposition $A = USV^T$ to obtain the shortest, least-squares solution to the matrix problem $Ax \neq b$. Assume that matrices U, S, and V are already available. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.
5. What is the Moore-Penrose Pseudo inverse? Explain how to obtain this pseudo-inverse using the singular value decomposition $A = USV^T$. Assume that matrices U, S, and V are already available. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.
6. Explain how a correlation filter works and how singular value decomposition can be used to filter data using this concept. Also, explain how to extract the filtered data using singular decomposition. Remember that stating facts is not enough; your answer must follow clear and logical steps (explain why and how for each step) supported by fundamental concepts.