Solution to HW 4

Name _____________________________

- On each item, grading is +, √ or 0 (full points, 1/2 points, 0 points)

<table>
<thead>
<tr>
<th>Problem #</th>
<th>Parts</th>
<th>Points Each Part</th>
<th>Your Part Points</th>
<th>Your Total Points</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. A</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. A.1</td>
<td>1</td>
<td>2</td>
<td>-1 each incorrect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.2</td>
<td>1</td>
<td>4</td>
<td>-1 each incorrect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.3</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B.1</td>
<td>1</td>
<td>2</td>
<td>-1 each incorrect</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C.1</td>
<td>1</td>
<td>2</td>
<td>-1 each incorrect</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C.2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C.3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>9</td>
<td>2</td>
<td></td>
<td>18+2</td>
<td>2 points extra credit</td>
</tr>
<tr>
<td>5.</td>
<td>4</td>
<td>2 total</td>
<td></td>
<td>-2</td>
<td>-1 each wrong</td>
</tr>
<tr>
<td>6. A</td>
<td>12</td>
<td></td>
<td></td>
<td>-4</td>
<td>-1 each incorrect</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2</td>
<td>1</td>
<td>/2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>1</td>
<td>2</td>
<td>/2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>1</td>
<td>2</td>
<td>/2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>1</td>
<td>2</td>
<td>/2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>1</td>
<td>2</td>
<td>/2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>1</td>
<td>2</td>
<td>/2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>2</td>
<td>2</td>
<td>/4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>4</td>
<td>1</td>
<td>/4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>2</td>
<td>2</td>
<td>/4</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td>/10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td>/10</td>
<td>Extra Credit</td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td></td>
<td></td>
<td>/100 + 10 E.C.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Gram-Schmidt Ortho-normalization

A) Write a Matlab program to execute Gram-Schmidt ortho-normalization.

Answer: Example program (any correct program valid for full credit)

```matlab
function [Vgs, ColNorms, Threshold] = GramSchmidt(V)
    % function [Vgs, ColNorms, Threshold] = GramSchmidt(V)
    % Return Gram Schmidt Orthogonlalization of vector set V.
    %
    % Inputs:
    % V: n x p array.
    %
    % Outputs:
    % Vgs: n x r matrix, the columns are ortho-normal basis vectors
    % ColNorms: ||vperp||
    % Threshold: Threshold used to determine which vectors are included in Vgs
    % [nVecLength, nCols] = size(V);
    % % This calculation determines a threshold
    % % Examples have been found where round off error exceed threshold without
    % % MaxVNorm = max(size(V)) * sqrt(max(diag(V'*V)));
    % % Threshold: Threshold used to determine which vectors are included in Vgs
    % Threshold = 10 * max(size(V)) * MaxVNorm * eps;
    % Vgs = []; jVecs_Done = 0; ColNorms = zeros(1, nCols);
    % for ii = 1:nCols,
    %     v_ortho = V(:,ii); % next vector to GS orthonormalize
    %     for jj = 1:jVecs_Done,
    %         v_ortho = v_ortho - Vgs(:,jj) * Vgs(:,jj)' * v;
    %     end
    %     vperp = v; % Projection of v onto ortho-comp. of Vgs
    %     ColNorms(ii) = norm(vperp);
    %     if ColNorms(ii) > Threshold, % If v is reasonable size
    %         if ColNorms(ii) < 1000*Threshold,
    %             warning('Sensitive result, vector norm within 1000x of threshold');
    %         else
    %             jVecs_Done = jVecs_Done + 1;
    %             Vgs(:,jVecs_Done) = vperp/norm(vperp);
    %         end
    %     end
    % end
    % return % GramSchmidt
```

Extra Credit

Total: /100 + 10 E.C.
B) Demonstrate that your Gram-Schmidt program reproduces the example data.

Answer:

\[ A = \begin{bmatrix} 2 & 3 & 5 \\ 5 & 6 & 11 \\ -1 & 2 & 1 \\ 9 & 1 & 10 \end{bmatrix} \]

\[ U = \text{GramSchmidt}(A) \]

\[ U = \begin{bmatrix} 0.1898 & 0.3854 & 0.4746 & 0.7036 \\ -0.0949 & 0.4135 & 0.8542 & -0.4306 \end{bmatrix} \]

C) What is the dimension of the space spanned by the set of vectors \( A \) of the example?

Answer: 2

2 Projection onto an ortho-normal basis

A) How can the equation

\[ b = \left( U^T U \right)^{-1} U^T y \]

be simplified when \( U \) is an ortho-normal basis, not just a basis?

Answer: The \( \left( U^T U \right)^{-1} \) term drops out, leaving \( b = U^T y \).

---

B) Given the vector \( y = \begin{bmatrix} 4 & 3 & 2 & 1 \end{bmatrix}^T \) and vector space \( U \) spanned by the columns of \( A \), above,

B.1) What is the projection of \( y \) onto \( U \)?

\[ \text{Py} = U^* b = U^* U^T y \]

\[ \text{Py} = \begin{bmatrix} 2.1008 \\ 1.4037 \\ 0.6890 \end{bmatrix} \]

B.2) Does \( y \) lie in space \( U \)?

If \( y \) lies in \( U \), then \( \text{Py} \) will equal \( y \). \( \text{Py} \) given above does not equal \( y \).

Answer: No

B.3) What component of \( y \) lies in \( U^\perp \)?

By the projection theorem, \( y \) can be split into \( \text{Py} \in U \) and \( y^\perp \in U^\perp \) with \( y^\perp = y - \text{Py} \)

Answer:

\[ \text{w} = y - U^* U^T y \]

\[ \text{w} = \begin{bmatrix} 1.8992 \\ -1.2002 \\ 0.5963 \\ 0.3110 \end{bmatrix} \]

B.4) What is the 2-norm of the component of \( y \) which lies in \( U \)?

Answer:

\[ \text{norm}( \text{Py} ) = 4.9498 \]

B.5) What is the 2-norm of the component of \( y \) which lies in \( U^\perp \)?

Answer:

\[ \text{norm}( \text{w} ) = 2.3452 \]
B.6) How are the answers to (4) and (5) related to $|y|^2$?

**Answer:**

$$|y|^2 = |Py|^2 + |y^\perp|^2$$

$$\text{norm}(y)^2 = 30, \quad \text{norm}(Py)^2 + \text{norm}(w)^2 = 30.0000$$

3 Four fundamental spaces of $A$

A) Given matrix $A$

$$A = \begin{bmatrix} 2 & 3 & 5 \\ 5 & 6 & 11 \\ -1 & 2 & 1 \\ 9 & 1 & 10 \end{bmatrix}$$

A.1) What are the names of the four fundamental spaces of $A$?

**Answer:**

- Row Space
- Column Space
- Null Space
- Left-Null Space

A.2) Using matrix $A$ and the rank of matrix $A$, what is the dimension of each of the fundamental spaces?

**Answer:**

- Row Space: 2
- Column Space: 2
- Null Space: 1
- Left-Null Space: 2

B) A matrix $B$ is 5x4, and has rank 3. What is the dimension of each of the fundamental spaces?

**Answer:**

- Row Space: 3
- Column Space: 3
- Null Space: 1
- Left-Null Space: 2
C) In a previous homework assignment, three teams collected experimental data to determine 3 parameters. Team 1 had a constrained case, team 2 an over-constrained case and team 3 an under-constrained case.

C.1) Make a table showing the dimension of the four fundamental spaces in each case.

<table>
<thead>
<tr>
<th></th>
<th>Team 1</th>
<th>Team 2</th>
<th>Team 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Column</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Null</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Left-Null</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Sizes of four fundamental spaces of Team problem from HW 1.

C.2) Which team(s) (or no team) can exactly fit the data, for any possible set of measurements?

Answer: Teams 1 and 3, in these cases there is no left-null space.

C.3) Which team(s) (or no team) has sets of parameters, \(|b| \neq 0\) such that \(Y = 0\)?

Answer: Team 3, in this case there is a null space.

4 GS Ortho-normalization in \(R^8\)

A) Form an orthonormal basis for the space spanned by the columns of \(W\).

Answer: there are many possible bases for any vector space. The one given by the basic Gram Schmidt algorithm is:

\[
U = \text{GramSchmidt}(W)
\]

\[
U = \begin{bmatrix}
-0.1927 & -0.2349 & 0.3448 & -0.2856 \\
0.3853 & -0.8731 & 0.0313 & 0.2248 \\
0.4335 & 0.1928 & 0.1335 & -0.3861 \\
0.2408 & -0.0854 & 0.4281 & -0.4036 \\
-0.2890 & -0.0708 & 0.0999 & 0.4641 \\
-0.1445 & -0.0354 & -0.0752 & -0.1314 \\
0.6744 & 0.3239 & -0.0530 & 0.4094 \\
-0.0963 & 0.1641 & 0.8188 & 0.3945 \\
\end{bmatrix}
\]

B) What is the dimension of the subspace spanned by \(W\)?

Answer: 4

C) Compute \(P_z\), the projection of the vector \(z = \begin{bmatrix} 1 & 2 & 3 & -1 & -2 & -3 & 0 & 1 \end{bmatrix}^T\) onto the subspace spanned by \(W\).

C.1) First try to form this projection without using the basis vectors given by the Gram-Schmidt algorithm. (5 points extra credit for finding the projection without using the ortho-normal basis vectors)

Answer: Here is one possible answer (all should come to \(P_z\))

\[
\begin{bmatrix}
-3.3824 \\
0.3824 \\
-0.1862 \\
-0.2305 \\
-0.1445 \\
0.4094 \\
0.8188 \\
0.3945
\end{bmatrix}
\]

\[
>> z = \begin{bmatrix} 1 & 2 & 3 & -1 & -2 & -3 & 0 & 1 \end{bmatrix}^T
\]

\[
>> P_z = z \cdot (W \cdot W^T)^{-1} \cdot W^T
\]
C.2) Find the projection using the ortho-normal basis vectors.

**Answer:**

```
>> z = [1 2 3 -1 -2 -3 0 1]’ ; Pz = U*U’*z
Pz = [ 0.4136
       1.6541
       1.4016
       1.5882
      -0.9988
      -0.3464
       1.0549
       0.4670 ]
```

D) What is the length of $Pz$?

**Answer:** $||Pz|| = 3.1373$

E) Define the angle between a vector and a subspace to be the smallest angle between the vector and any vector in the subspace.

What is the angle between vector $z$ and subspace $W$?

**Answer:** The vector lying in $W$ which has the smallest angle to $z$ is $Pz$.

Evaluating that angle

```
>> DirCosine = z’*Pz / (norm(z)*norm(Pz))
DirCosine = 0.5826
>> DirCosineAngle = acosd(DirCosine)
DirCosineAngle = 54.3681
```

F) Compute $P^\perp z$ the projection of $z$ onto subspace $W^\perp$.

**Answer:** $P^\perp z = z - Pz$

```
>> zperp = z - Pz
zperp = [ 0.5864
          0.3459
         1.5984
        -2.5882
       -2.6536
       -1.0549
        0.5330 ]
```
G) What is the angle between vector $P^\perp z$ and subspace $W$?

Answer: This angle must be 90 degrees. This can be shown by showing that $P^\perp z$ is orthogonal to each of the basis vectors, therefore it will be orthogonal to any linear combination of basis vectors, and thus to any vector in the space.

```matlab
>> zperp'*U
ans =
    1.0e-15 *
   -0.4441   -0.1665   -0.2220    0.2220
```

H) Writing $W = \begin{bmatrix} w_1 & w_2 & \ldots & w_p \end{bmatrix}$, what is the length of the longest vector $w_j$?

```matlab
>> for jj = 1:size(W,2), NW(jj) = norm(W(:,jj)); end
NW = [2.0761 2.4495 2.6287 3.2848 3.6042 1.9416 4.0472 4.3046]
>> max(NW) = 4.3046
```

Answer: The length of the longest $w_j$ is 4.3046

I) Writing the orthonormal basis for $W$ as $V = \begin{bmatrix} v_1 & v_2 & \ldots & v_r \end{bmatrix}$, what is the length of the longest vector $v_j$?

Answer: 1.0

5 Fundamental spaces of matrix $W$

- What are the dimensions of the 4 fundamental spaces of matrix $W$ of problem 4.

Answer: Row Space: 4, Column Space: 4, Null Space: 5, Left-Null Space: 4

6 The projection theorem and four fundamental spaces applied to $Y = A b$

A) Writing $Y = A b$, so that $A$ is the regressor matrix, what is the dimension of the input and output space and each of the four fundamental spaces of the regressor matrix (and we may say of the estimation problem) in each case.

Answer:

<table>
<thead>
<tr>
<th>Team 2</th>
<th>n</th>
<th>p</th>
<th>row</th>
<th>column</th>
<th>null</th>
<th>left null</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

B) In each of the two cases, what rank must the regressor matrix have to be called “full rank”.

Answer: Team 2: rank 3, Team 3: rank 2.

C) When we estimate the parameters, $\hat{b}$, we can write

$$\hat{y} = A \hat{b}$$

$$\tilde{y} = \hat{y} - \bar{y}$$

where $\bar{y}$ is the measured data.

Considering $R^n$ and $R^p$ and the four fundamental spaces, what space (or spaces) contain each of $\bar{y}$, $\tilde{y}$ and $\hat{y}$.

Answer:

$$y \in R^n, \quad \tilde{y} \in \text{col}A, \quad \hat{y} \in \text{null}A$$
**Additional Discussion:** $\mathbf{y}$ exists in $\mathbb{R}^n$. Vector $\hat{\mathbf{y}}$ is the projection of $\mathbf{y}$ onto the column space of $A$, and $\tilde{\mathbf{y}}$ is the projection of $\mathbf{y}$ onto the left-null space. The projection theorem tells us that there is a $\hat{\mathbf{y}} \in \text{col}(A)$ which minimizes $||\mathbf{y} - \hat{\mathbf{y}}||$, and that with $\tilde{\mathbf{y}} = - (\mathbf{y} - \hat{\mathbf{y}})$, then $\tilde{\mathbf{y}} \perp \text{col}(A)$.

**D)** (Essay Question) Building on your response to part C, describe a way we can determine $\hat{\mathbf{y}}$ using the Gram-Schmidt algorithm and projection, without ever solving a matrix inverse or solving for $\hat{\mathbf{b}}$. Include the equation(s) giving $\hat{\mathbf{y}}$ with definitions for all terms.

**Answer:** Since we know that $\hat{\mathbf{y}}$ is the projection of $\mathbf{y}$ onto the column space of $A$, $\hat{\mathbf{y}}$ can be found by projecting $\mathbf{y}$ onto a basis for the column space for $A$. This projection is given by:

$$
\hat{\mathbf{y}} = V_c V_c^T \mathbf{y} 
$$

where $V_c$ is a collection of ortho-normal basis vectors for $\text{col}A$. Neither Eqn (1) nor applying the Gram-Schmidt algorithm to $A$ involves forming a matrix inverse.

**E)** Using your result in part D, compute $\hat{\mathbf{y}}$ for the data of the second team, without ever solving a matrix inverse or solving for $\hat{\mathbf{b}}$.

```matlab
>> Vcol = GramSchmidt(Phi2)
Vcol =
 0.4082 -0.5976 0.4616
 0.4082 -0.3586 0.0438
 0.4082 -0.1195 -0.4527
 0.4082 0.1195 -0.5279
 0.4082 0.3586 -0.0694
 0.4082 0.5976 0.5446

>> Yhat2 = Vcol*Vcol'*Ybar2
Yhat2 =
 1.5872
 -0.8325
-3.4075
-5.1494
-5.8357
-6.2145

• Double check by computing $\hat{\mathbf{y}}$ the usual way (not required by the problem, but double checking is always a good idea).

```matlab
>> b2hat = (Phi2'*Phi2)Phi2'*Ybar2
b2hat =
 1.9984
 -1.4999
 1.1973
 >> Yhat2b = Phi2*b2hat
Yhat2b =
 1.5872
 -0.8325
-3.4075
-5.1494
-5.8357
-6.2145
```
F) (Essay Question) Building on your response to part (C), describe a way we can determine $\tilde{y}$ using the Gram-Schmidt algorithm and projection, without ever solving a matrix inverse or solving for $\hat{b}$ or $\hat{y}$.

**Answer:** $\tilde{y}$ can be found by projecting $y$ onto a basis for the left null space for $A$. This is given by:

$$\tilde{y} = V_{ln} V_{ln}^T y$$

(2)

where $V_{ln}$ is a collection of orthonormal basis vectors for the left null space of $A$.

(We may also write

$$\tilde{y} = \left( I - V_{c} V_{c}^T \right) y$$

(3)

G) Using your result in part (F), compute $\tilde{y}$ for the data of the second team, without ever solving a matrix inverse or solving for $\hat{b}$ or $\hat{y}$.

```
>> Vln = GramSchmidt(eye(6)-Vcol*Vcol')
Vln =
 0.5129  -0.0000  -0.0000
-0.7821   0.3019  -0.0000
-0.0568  -0.7754   0.0980
 0.2894   0.4162  -0.5323
 0.1553   0.2862  -0.7707
-0.1188  -0.2289  -0.3363
>> Ytilde2 = Vln*Vln'*Ybar2
Ytilde2 =
 0.0063
-0.0102
 0.0065
-0.0102
 0.0003
-0.0013
```

H) Power

In a qualitative way we can speak of the “power” in a signal, even if that signal is not specifically related to watts. In this sense the “power” is given by the sum-squared value of the signal. For example, for team 2’s data the “power” in $\tilde{y}$ can be given by

$$P_{\tilde{y}} = \sum_{k=1}^{6} \tilde{y}_k^2$$

H.1) For team 2, what is the power in $y$, $\hat{y}$ and $\tilde{y}$.

```
>> PYbar2 = sum(Ybar2.^2)
PYbar2 = 114.015
>> PYhat2 = sum(Yhat2.^2)
PYhat2 = 114.015
>> PYtilde2 = sum(Ytilde2.^2)
PYtilde2 = 0.0002
```

(Note: PYbar2 and PYhat2 come out looking the same because the values are printed to only 3 places)

H.2) How do these values relate to the lengths of the vectors.

**A:** The “power” values are the $L_2$ norms squared of the vectors.

```
>> norm(Ybar2)^2
ans = 114.015
>> norm(Yhat2)^2
ans = 114.015
>> norm(Ytilde2)^2
ans = 0.00019
```
I) Power per degree of freedom

I.1) What is the power per degree of freedom in $\hat{y}$, $\tilde{y}$.

**Answer:** For team 2 there are 3 degrees of freedom in each of $\hat{y}$ and $\tilde{y}$,

```matlab
>> PpDOF_Yhat2 = (1/3)*sum(Yhat2.^2)
PpDOF_Yhat2 = 38.005
>> PpDOF_Ytilde2 = (1/3)*sum(Ytilde2.^2)
PpDOF_Ytilde2 = 0.000063
```

I.2) What is the ratio of power per degree of freedom in $\hat{y}$ and the power per degree of freedom in $\tilde{y}$?

```matlab
>> F = PpDOF_Yhat2 / PpDOF_Ytilde2
F = 598851.30
```

I.3) (Essay Question) If this ratio were small, what would we think about the quality of our determined parameters? (This ratio is called the “Fisher statistic”.)

**Answer:** If the ratio of power per degree of freedom fit by the model were small, we would conclude that the model is mainly fitting noise in the data.

I.4) The dimensions of what two fundamental spaces play a role in determining the Fisher statistic

**Answer:** The column space and the left-null space. $\dim \text{col}(A)$ is the denominator for determining the power per DOF of $\hat{y}$, and $\dim \text{null}(A)$ is the denominator for determining the power per DOF of $\tilde{y}$.

J) The third team has 3 parameters and only 2 data points, which means, among other things, that they can exactly fit the model to the data. In fact, many possible values for the parameters exactly fit the data.

J.1) Using in some way at least one of the four fundamental spaces, write an expression for all possible sets of parameters that fit the model to team 3’s data.

**Answer:** Given a particular solution to

$$ Y = Ab $$ (4)

call this $b_p$, the set of all solutions is given by:

$$ b_2 = b_p + b_n $$

were $b_n \in \text{null} A$.

**First,** find the particular solution

```matlab
>> b3p = Phi3'*inv(Phi3*Phi3')*Ybar3b3p = 0.7411 -0.9224 1.9517
```

Now, find a basis vector for the null space

```matlab
>> Vr3 = GramSchmidt(Phi3') % Basis for the row sp.>> Vn3 = GramSchmidt(eye(3)-Vr3*Vr3') % Basis for the null sp.Vn3 = 0.7993 -0.3660 -0.4765
```

Homework Solution 4 (Revised: Oct 04, 2018) Page 17
So the set of all exact solutions is given as:

\[ \mathbf{b} = \mathbf{b}_p + a_1 \mathbf{b}_n = \begin{bmatrix} 0.7411 \\ -0.9224 \\ 1.9517 \end{bmatrix} + a_1 \begin{bmatrix} 0.7993 \\ -0.3660 \\ -0.4765 \end{bmatrix} \] (5)

J.2) Using your result for part (1), find a solution for the parameters that exactly fits the data and also satisfies \( b_2 = 0 \).

\[ \begin{align*}
&\text{Answer: solve for } a_1 \text{ in Eqn (5) to give } b_2 = 0 \\
&\text{• Determine a contribution from the null space that make } b_2 = 0.
\end{align*} \]

\[ \begin{align*}
&\text{• Determine the solution with this contribution from the null space}
\end{align*} \]

\[ \begin{align*}
&\text{• Verify that } \mathbf{b} \text{ satisfies Eqn (??)}
\end{align*} \]

\[ \begin{align*}
&\text{Proof by Contradiction: For a proof by contradiction, assume that } \mathbf{x} \neq \mathbf{y} \text{ and show a contradiction. Assuming that } \mathbf{x} \neq \mathbf{y}, \text{ then there exists a non-zero vector } \mathbf{w} = \mathbf{x} - \mathbf{y}. \text{ Choose } \mathbf{z} = \mathbf{w} = \mathbf{x} - \mathbf{y}, \text{ also note } \mathbf{x} = \mathbf{y} + \mathbf{w}, \text{ then}
\end{align*} \]

\[ \begin{align*}
&\langle \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{w} \rangle \\
&\langle \mathbf{x}, \mathbf{z} \rangle = \langle (\mathbf{y} + \mathbf{w}), \mathbf{w} \rangle = \langle \mathbf{y}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\
\end{align*} \]

so \( \langle \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{z} \rangle + ||\mathbf{w}||^2 \). Since \( ||\mathbf{w}|| \neq 0 \), it follows that \( \langle \mathbf{x}, \mathbf{z} \rangle \neq \langle \mathbf{y}, \mathbf{z} \rangle \), but this contradicts the hypothesis that \( \langle \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{z} \rangle \) for every \( \mathbf{z} \in V \).

\[ \text{QED (1)} \]
Example direct proof (using an ortho-normal basis): Let \( \{v_1, ..., v_n\} \) be an ortho-normal basis for \( V \); and recall that a representation for any vector \( x \in V \) is
\[
x = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n
\]
And let the representation for \( y \) be
\[
y = \gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n
\]
If \( \langle x, z \rangle = \langle y, z \rangle \forall z \in V \), then \( \langle x, z \rangle = \langle y, z \rangle \) for any set \( z \in \{z_1, ..., z_p\} \) that we choose.

Choose \( z_j \) to be the basis vectors of \( V \), \( z_1 = v_1 \), etc. Since \( \{z_1, ..., z_n\} \) forms an orthonormal basis, it follows that the basis coefficients are given by:
\[
\langle x, z_j \rangle = \beta_j \quad \text{and} \quad \langle y, z_j \rangle = \gamma_j.
\]
The hypothesis that \( \langle x, z \rangle = \langle y, z \rangle \) implies that \( \beta_1 = \gamma_1, \beta_2 = \gamma_2, \) etc. Since the basis coefficients of \( x \) are the same as the basis coefficients of \( y \), the representations of the two vectors are the same, which proves that \( x = y \).

QED (2)

8 Text problem 2.12, (Extra credit 10 Points)

Question: If it is known that a valid basis for the space of polynomials of degree less than or equal to three is \( \{1, t, t^2, t^3\} \), show that every polynomial has a unique representation as
\[
p(t) = a_3 t^3 + a_2 t^2 (1 - t) + a_1 t (1 - t)^2 + a_0 (1 - t)^3
\]

• Hint: Think of the transformation from

Representation on \( B \) basis: \( p(t) = b_3 t^3 + b_2 t^2 + b_1 t + b_0 \)
to

Representation on \( A \) basis: \( p(t) = a_3 t^3 + a_2 t^2 (1 - t) + a_1 t (1 - t)^2 + a_0 (1 - t)^3 \)
as a change of basis. Finding the transformation matrix to transform from one basis to the other will permit you to show that every polynomial has a unique representation on the \( A \) basis.

Answer: The most straight forward way to show (prove) the property described in the question is to show the method of calculating \[
\begin{bmatrix}
a_0 & a_1 & a_2 & a_3 \\
b_0 & b_1 & b_2 & b_3
\end{bmatrix}
\]
given

Definitions:

\( A(t) = \begin{bmatrix} (1 - t)^3, & t (1 - t)^2, & t^2 (1 - t), & t^3 \end{bmatrix} \)
Now
\[ p(t) = A(t) \quad x = \begin{bmatrix}
(1-t)^3, & t (1-t)^2 & t^2 (1-t) & t^3
\end{bmatrix} \begin{bmatrix}
a_0 \\
a_1 \\
a_2 \\
a_3
\end{bmatrix}, \quad \text{where} \quad x = \begin{bmatrix}
a_0 \\
a_1 \\
a_2 \\
a_3
\end{bmatrix}
\]

**B-basis:** The set of basis functions for the \( B(t) \) basis is:
\[ B(t) = \begin{bmatrix}
1, & t, & t^2, & t^3
\end{bmatrix}
\]

Now
\[ p(t) = B(t) \quad x = \begin{bmatrix}
b_0 \\
b_1 \\
b_2 \\
b_3
\end{bmatrix}, \quad \text{where} \quad x = \begin{bmatrix}
b_0 \\
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\]

**Solution:** If we can find \( B_A \), the transformation from \( A(t) \) basis to \( B(t) \) basis, such that
\[ A(t) = B(t) \quad B_A \]

then
\[ p(t) = A(t) \quad x = B(t) \quad B_A \quad x = B(t) \quad x \]

which shows that
\[ x = B_A \quad x \]

- If \( B_A \) is invertible, \( x = (B_A)^{-1} \quad x = B_A \quad x \).

**From Eqn (7), each of the functions in \( A(t) \) must be given by**
\[ a_i(t) = B(t) \quad v_i \]

where \( v_i \) is a column of \( B_A \).

\[ a_0(t) = (1-t)^3 = 1 - 3t + 3t^2 - t^3 = B(t) \]
\[ a_1(t) = t (1-t)^2 = t - 2t^2 + t^3 = B(t) \]
\[ a_2(t) = t^2 (1-t) = t^2 - t^3 = B(t) \]
\[ a_3(t) = t^3 = B(t) \]
So

\[ A(t) = B(t) \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ -1 & 1 & 1 & 1 \end{bmatrix} = B(t) B_A^T \]

- Now

\[ B_x = B_A^T A_x \]

And if \( B_A^T \) is invertible,

\[ A_x = B_A^{-1} B_x \]

Which shows that every \( B_x \) corresponds to a unique \( A_x \), and thus that every polynomial in \( B \) has a unique representation in \( A \).

**Note on connection to Bay, section 2.2.4**

- In Bay, section 2.2.4, Bay introduces geometric interpretation of a change of basis. He develops

\[
\begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_r \end{bmatrix} = \begin{bmatrix} | & | & | \\ \mathbf{\hat{v}}_1 & \mathbf{\hat{v}}_2 & \cdots & \mathbf{\hat{v}}_r \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rr} \end{bmatrix} \begin{bmatrix} | & | & | \\ \mathbf{\hat{v}}_1^T & \mathbf{\hat{v}}_2^T & \cdots & \mathbf{\hat{v}}_r^T \end{bmatrix}
\]

- So the transformation from \( \mathbf{v} \) to \( \mathbf{\hat{v}} \) is given by the representation of the \( \mathbf{v} \) basis vectors on \( \mathbf{\hat{v}} \).

- \( B_A^T \) is given by the representation of the \( A(t) \) basis functions on \( B(t) \).