Assignment 2 - Vector Spaces

1 Skill Builder

Sessions 4.1 & 4.2: The Geometry of Vectors

1. Prove the following distributive properties hold for vectors $\vec{u}$, $\vec{v}$, $\vec{w}$ in $\mathbb{R}^2$ and scalars c and d.

   (a) $c(\vec{u}+\vec{v}) = c\vec{u} + c\vec{v}$

   (b) $(c+d)\vec{v} = c\vec{v} + d\vec{v}$

   (c) $\vec{u} \cdot (\vec{v}+\vec{w}) = \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}$

   (d) (Optional) Generalize one (or more) of your proofs to vectors in $\mathbb{R}^n$.

(a) $c(\vec{u}+\vec{v})$=
$= c(\langle u_1, u_2, ..., u_n\rangle, \langle v_1, v_2, ..., v_n\rangle) =$
$= c(\langle u_1 + v_1, u_1 + v_1, ...u_n + v_n\rangle)$
$= \langle c(u_1 + v_1), c(u_2 + v_2), ..., c(u_n + v_n)\rangle =$
$= \langle cu_1 + cv_1, cu_2 + cv_2, .., cv_n + cu_n\rangle =$
$= c(\langle u_1, u_2, ..., u_n\rangle) + c(\langle v_1, v_2, ..., v_n\rangle) =$
$= c\vec{u} + c\vec{v}$

(b) $(c+d)\vec{v}$=
$= (c+d)\langle v_1, v_2, ..., v_n\rangle =$
$= \langle (c+d)v_1, (c+d)v_2, ..., (c+d)v_n\rangle =$
$= \langle cv_1 + dv_1, cv_2, +dv_2, ..., cv_n + dv_n\rangle =$
$= c\langle v_1, v_2, ..., v_n\rangle + d\langle v_1, v_2, ..., v_n\rangle =$
$= c\vec{v} + d\vec{v}$

(c) $\vec{u}\cdot(\vec{v}+\vec{w}) =$
$= \langle u_1, u_2, ..., u_n\rangle \cdot \langle v_1 + w_1, v_2 + w_2, ..., v_n + w_n\rangle =$
$= \langle u_1v_1 + u_1w_1 + u_2v_2 + u_2w2 + ... + u_nv_+u_nw_n\rangle =$
$= \langle u_1v_1 + u_2v_2 + ...u_nw_n\rangle + \langle u_1w_1 + u_2w_2 + ... + u_nw_n\rangle =$
$= \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}$

2. Consider the following system of equations:

$$\begin{aligned} x_1 + x_2 - 2x_3 + x_4 &= 1 \\ x_1 + 4x_2 - 4x_3 + 6x_4 &= 5 \\ 3x_1 + 6x_2 - 8x_3 + 8x_4 &= 7 \end{aligned}$$

Find the parametric solution of this system. Decompose the solution into column vector form. Describe the geometry of the solution set. (Hint: Where does the solution live? What is the dimension of the solution space?)

Represent the equation as a matrix:

$$A = \begin{bmatrix} 1 & 1 & -2 & 1 & | & 1 \\ 1 & 4 & -4 & 6 & | & 5 \\ 3 & 6 & -8 & 8 & | & 7 \end{bmatrix}$$

Turn the matrix into RREF:

$$A = \begin{bmatrix} 1 & 0 & -4/3 & -2/3 & | & -1/3 \\ 0 & 1 & -2/3 & 5/3 & | & 4/3 \\ 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}$$

Write the parametric form:

$x_1 + 0x_2 - 4/3x_3 - 2/3x_4 = -1/3$
$0x_1 + x_2 - 2/3x_3 + 5/3x_4 = 4/3$
$0x_1 + 0x_2 - 0x_3 + 0x_4 = 0$

Substitute the parametric columns with t and z, respectively:
$x_3 = t$
$x_4 = z$
Represent the final equation in column vectors:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1/3 + 4/3 * t + 2/3 * z \\ 4/3 + 2/3 * t - 5/3 * z \\ t \\ z \end{bmatrix} = \begin{bmatrix} -1/3 \\ 4/3 \\ 0 \\ 0 \end{bmatrix} + \text{t} \begin{bmatrix} 4/3 \\ 2/3 \\ 1 \\ 0 \end{bmatrix} + \text{z} \begin{bmatrix} 2/3 \\ -5/3 \\ 0 \\ 1 \end{bmatrix}$$

There are two parametric columns and two pivot ones. The system is consistent, yet it has infinitely many solutions. The solution lies in a 3D space and it is a hyperplane because of the two free variables.

Sessions 5.1, 5.2 & 6.2: Vector Spaces

1. Show that the set of 3×3 lower triangular matrices T is a vector space under matrix addition and scalar multiplication.

T= $\begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix}$

   (a) Closure under addition
   $$\begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix} + \begin{bmatrix} x & 0 & 0 \\ y & v & 0 \\ z & u & t \end{bmatrix} = \begin{bmatrix} a+x & 0 & 0 \\ b+y & c+v & 0 \\ d+z & e+u & f+t \end{bmatrix}$$
   The sum is a lower triangular matrix.

   (b) Closure under scalar multiplication
   $$\alpha \begin{bmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{bmatrix} = \begin{bmatrix} \alpha a & 0 & 0 \\ \alpha b & \alpha c & 0 \\ \alpha d & \alpha e & \alpha f \end{bmatrix}$$
   The product is a lower triangular matrix.

2. Show that the set of 4×4 diagonal matrices D is a subspace of $M_{4\times 4}$.

D = $\begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}$

   (a) Closure under addition
   $$\begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix} + \begin{bmatrix} e & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & g & 0 \\ 0 & 0 & 0 & h \end{bmatrix} = \begin{bmatrix} a+e & 0 & 0 & 0 \\ 0 & b+f & 0 & 0 \\ 0 & 0 & c+g & 0 \\ 0 & 0 & 0 & d+h \end{bmatrix}$$
   The sum is a $M_{4\times 4}$ matrix.

(b) Closure under scalar multiplication

$$\alpha \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix} = \begin{bmatrix} \alpha a & 0 & 0 & 0 \\ 0 & \alpha b & 0 & 0 \\ 0 & 0 & \alpha c & 0 \\ 0 & 0 & 0 & \alpha d \end{bmatrix}$$

The sum is a $M_{4\times 4}$ matrix.

(c) Zero vector

$$\begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}$$

The sum between the zero vector and the matrix is the $M_{4\times 4}$ matrix itself.

3. For each of the following sets of vectors S in a vector space V: (i) Describe the subspace spanned by the set S. (ii) Determine the dimension of S. (iii) Extend the set S to a basis for V.

   (a) $S = \{\langle 1, 0, -2, 1\rangle, \langle 0, 1, 0, 1\rangle\}$ in $\mathbb{R}^4$

      i. Describe the subspace spanned by the set S

      $$\begin{aligned} \text{span}(S) &= \{a\langle 1, 0, -2, 1\rangle + b\langle 0, 1, 0, 1\rangle\} = \\ &= \{\langle a, 0, -2a, a\rangle + \langle 0, b, 0, b\rangle\} = \\ &= \{\langle a, b, -2a, a+b\rangle\} \end{aligned}$$

      ii. Determine the dimension of S

      $\dim(S) = 4$

      iii. Extend the set S to a basis for V.
      Considering that the dimension of $\mathbb{R}^4$ is 4, then we need two more vectors to form a basis. In that case, we could add two standard unit vectors to the original ones and check for their linear independence, as the span will not be affected by adding two standard unit vectors.

      $$\text{S}= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

$$S_{rref} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Since we have obtained the identity matrix, we prove that the combination of these vectors is linearly independent, thus we have made sure that these 4 vectors work as a basis for the vector space V.

Basis $(V) = \{\langle 1, 0, -2, 1\rangle, \langle 0, 1, 0, 1\rangle, \langle 0, 0, 1, 0\rangle, \langle 0, 0, 0, 1\rangle\}$

(b) $S = \{x^2 - 4, 3x, 2x^2 - 1\}$ in $P(3)$ (i.e., the space of polynomials of degree at most 3)

i. Describe the subspace spanned by the set S

$\text{span}(S) = \{a(x^2 - 4) + b(3x) + c(2x^2 - 1)\} =$
$= \{ax^2 - 4a + 3xb + 2x^2c - c\}$

ii. Determine the dimension of S

dim $(S) = 3$

iii. Extend the set S to a basis for V.

Check for linear independence:

- Represent the equations in terms of x:

$a(x^2 - 4) + b(3x) + c(2x^2 - 1) = 0x^2 + 0x + 0$

$x^2(a + 2c) + x(b) + 1(4a - c) = 0$

$a + 2c = 0$
$b = 0$
$-4a - c = 0$

- Create the matrix:

$$\begin{bmatrix} 1 & 0 & 2 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ -4 & 0 & -1 & | & 0 \end{bmatrix}$$

- Find RREF:

$$\begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}$$

The system has only one solution thus it is linearly independent. Thus $\{x^2, x, 1\}$ make a basis for the given vector, since they span the space and are linearly independent.

Sessions 7.1 & 7.2: Orthonormal Bases

1. Consider the vectors $v_1 = \langle 1, 1, 0\rangle$, $v_2 = \langle 0, 1, -4\rangle$, $v_3 = \langle -3, 0, 2\rangle$.

   (a) Show that $B = \{v_1, v_2, v_3\}$ is a basis of $\mathbb{R}^3$.

   (b) Use the Gram Schmidt Algorithm to transform B into an orthonormal basis.

   (c) Repeat the process, this time using a different ordering on the vectors of B (i.e., if you started the algorithm with v_1 before, start with v_3 this time.

   (d) Compare the results of your two applications of Gram Schmidt. What happened? What does it tell you about the algorithm?

(a) B is basis of $\mathbb{R}^3$.

Check for spanning:

$$\begin{aligned} &B = \{a_1 v_1 + a_2 v_2 + a_3 v_3\} = \\ &\{a_1\langle 1, 1, 0\rangle + a_2\langle 0, 1, -4\rangle + a_3\langle -3, 0, 2\rangle\} = \\ &= \{\langle a_1, a_1, 0\rangle + \langle 0, a_2, -4a_2\rangle + \langle -3a_3, 0, 2a_3\rangle\} = \\ &= \langle a_1 - 3a_3, a_1 + a_2, -4a_2 + 2a_3\rangle \end{aligned}$$

It spans $\mathbb{R}^3$.

Check for linear independence:

-Create the matrix:

$$\begin{bmatrix} 1 & 1 & 0 & | & 0 \\ 0 & 1 & -4 & | & 0 \\ -3 & 0 & 2 & | & 0 \end{bmatrix}$$

-Turn it to RREF:

$$\begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}$$

The system is linearly independent.

Thus, B forms a basis for $\mathbb{R}^3$.

(b) Gram-Schmidt Algorithm:

1. Normalize $\vec{v_1}$

$\vec{v_1} = \vec{u_1}$

$\vec{u_1} = \langle 1, 1, 0 \rangle$

$|\vec{u_1}| = \sqrt{1+1+0} = \sqrt{2}$

The normalized vector:

$\vec{e_1} = \langle 1, 1, 0 \rangle / \sqrt{2} = \langle 1/\sqrt{2}, 1/\sqrt{2}, 0 \rangle$

2. Normalize and orthogonalize $\vec{v_2}$

$\vec{u_2} = \vec{v_2} - proj_{u_1}(\vec{v_2})$

$\vec{u_2} = \vec{v_2} - \frac{\vec{u_1} \cdot \vec{v_2}}{|\vec{u_1}|^2)} * \vec{u_1} =$
$= \langle 0, 1, -4 \rangle - \frac{\langle 1,1,0 \rangle \cdot \langle 0,1,-4 \rangle}{2} * \langle 1, 1, 0 \rangle =$
$= \langle 0, 1, -4 \rangle - \frac{1}{2} * \langle 1, 1, 0 \rangle =$
$= \langle 0, 1, -4 \rangle - \langle \frac{1}{2}, \frac{1}{2}, 0 \rangle =$
$= \langle -\frac{1}{2}, \frac{1}{2}, -4 \rangle$

The normalized vector:

$\vec{e_2} = \langle -\frac{1}{2}, \frac{1}{2}, -4 \rangle \ \div \sqrt{(-\frac{1}{2})^2 + (\frac{1}{2})^2 + (-4)^2} =$
$= \langle -\frac{1}{2}, \frac{1}{2}, -4 \rangle \div \sqrt{\frac{33}{2}} =$

$= \langle -\frac{1}{66}, \frac{1}{66}, -\frac{4\sqrt{2}}{\sqrt{33}} \rangle$

3. Normalize and orthogonalize $\vec{v_3}$

$\vec{u_3} = \vec{v_3} - proj_{u_1}(\vec{v_3}) - proj_{u_2}(\vec{v_3})$

$proj_{u_1}(\vec{v_3}) = \langle \frac{-3}{2}, \frac{-3}{2}, 0 \rangle$

$proj_{u_2}(\vec{v_3}) = \langle \frac{13}{66}, \frac{-13}{66}, \frac{52}{33} \rangle$

$\vec{u_3} = \langle -3, 0, 2 \rangle - \langle \frac{-3}{2}, \frac{-3}{2}, 0 \rangle - \langle \frac{13}{66}, \frac{-13}{66}, \frac{52}{33} \rangle$
$= \langle \frac{-56}{33}, \frac{56}{33}, \frac{14}{33} \rangle$

The normalized vector:

$\vec{e_3} = \langle \frac{-56}{33}, \frac{56}{33}, \frac{14}{33} \rangle \div \frac{14}{\sqrt{33}} =$
$= \langle \frac{-4}{\sqrt{33}}, \frac{4}{\sqrt{33}}, \frac{1}{\sqrt{33}} \rangle$

The orthonormal basis is $\{\vec{e_1}, \vec{e_2}, \vec{e_3}\} = \{\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \rangle, \langle -\frac{1}{66}, \frac{1}{66}, -\frac{4\sqrt{2}}{\sqrt{33}} \rangle, \langle \frac{-4}{\sqrt{33}}, \frac{4}{\sqrt{33}}, \frac{1}{\sqrt{33}} \rangle$**.**

(c) Different order

Now we swap $\vec{v_1}$ with $\vec{v_2}$.

1. Normalize $\vec{v_2}$

$\vec{v_2} = \vec{u_2}$

$\vec{u_2} = \langle 0, 1, -4 \rangle$

$|\vec{u_2}| = \sqrt{17}$

The normalized vector:

$\vec{e_2} = \langle 0, 1, -4 \rangle / \sqrt{17} = \langle 0, 1/\sqrt{17}, -4/\sqrt{17} \rangle$

2. Normalize and orthogonalize $\vec{v_1}$

$\vec{u_1} = \vec{v_1} - proj_{u_2}(\vec{v_1})$

$proj_{u_2}\vec{v_1} = \langle 0, 1/17, -4/17\rangle$

$\vec{u_1} = \vec{v_1} - \frac{\vec{u_2}\cdot\vec{v_1}}{|\vec{u_2}|^1)} * \vec{u_2} =$
$= \langle 1, 1, 0\rangle - \langle 0, 1/17, -4/17\rangle = \langle 1, 16/17, 4/17\rangle$

The normalized vector:

$\vec{e_1} = \frac{\langle 1,16/7,4/7\rangle}{\sqrt{\frac{33}{17}}} = \langle\sqrt{\frac{17}{33}}, 16/\sqrt{561}, -4/\sqrt{561}\rangle$

3. Normalize and orthogonalize $\vec{v_3}$:

$\vec{u_3} = \vec{v_3} - proj_{u_2}(\vec{v_3}) - proj_{u_1}(\vec{v_3})$

$proj_{u_2}(\vec{v_3}) = \langle 0, \frac{-8}{17}, \frac{32}{17}\rangle$

$proj_{u_1}(\vec{v_3}) = \langle\frac{-43}{33}, \frac{-688}{561}, \frac{-172}{561}\rangle$

$\vec{u_3} = \langle -3, 0, 2\rangle - \langle 0, \frac{-8}{17}, \frac{32}{17}\rangle - \langle\frac{-43}{33}, \frac{-688}{561}, \frac{-172}{561}\rangle$
$= \langle\frac{-56}{33}, \frac{56}{33}, \frac{14}{33}\rangle$

The normalized vector:

$\vec{e_3} = \langle\frac{-56}{33}, \frac{56}{33}, \frac{14}{33}\rangle \div \frac{14}{\sqrt{33}} =$
$= \langle\frac{-4}{\sqrt{33}}, \frac{4}{\sqrt{33}}, \frac{1}{\sqrt{33}}\rangle$

The orthonormal basis is $\{\vec{e_2}, \vec{e_1}, \vec{e_3}\}$
$= \{\langle\sqrt{\frac{17}{33}}, 16/\sqrt{561}, -4/\sqrt{561}\rangle, \langle\sqrt{\frac{17}{33}}, 16/\sqrt{561}, -4/\sqrt{561}\rangle, \langle\frac{-4}{\sqrt{33}}, \frac{4}{\sqrt{33}}, \frac{1}{\sqrt{33}}\rangle$.

(d) Comparison

When we changed the vectors, we get another orthonormal basis. This can be explained by the fact that the second and the third vector are a projection of the first, thus their relationships are altered. However, since we only swapped $\vec{v_1}$ and $\vec{v_2}$, we find that the third vector in the basis remains the same. That is because the projection of $\vec{v_1}$ and $\vec{v_2}$ in $\vec{v_3}$ do not change by changing their order.

Session 8.1: Change of Basis

1. Let $B_1 = \{\langle 1,2,0\rangle, \langle 1,3,2\rangle, \langle 0,1,3\rangle\}$ and $B_2 = \{\langle 1,1,0\rangle, \langle 0,1,1\rangle, \langle 1,2,2\rangle\}$.

   (a) Show that both B_1 and B_2 form bases of $\mathbb{R}^3$.

   In order to form basis for $\mathbb{R}^3$, the vectors need to be linearly independent and span the vector space:

   Spanning:

   $\text{span}(B_1) = \{a\vec{v_1} + b\vec{v_2} + c\vec{v_3}\} =$
   $= \langle a_1, 2a_1, 0\rangle + \langle a_2, 3a_2, 2a_2\rangle + \langle 0, a_3, 3a_3\rangle = \langle a_1 + a_2, 2a_1 + 3a_2 + a_3, 2a_2 + 3a_3\rangle$
   ... it spans the $\mathbb{R}^3$.

   $\text{span}(B_2) = \{a\vec{v_1} + b\vec{v_2} + c\vec{v_3}\} =$
   $= \langle a_1 + a_3, a_1 + a_2 + a_3, a_2 + 2a_3\rangle$...it spans the $\mathbb{R}^3$.

   Linear Independence:

   $$\mathrm{B}_1 = \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 2 & 3 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$
   $$\mathrm{B}_2 = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

   Given their $RREF$, that have all pivot columns, both basis have linearly independent vectors.

   Since both vector span the space and are linearly independent, they make bases for $\mathbb{R}^3$.

   (b) Find the coordinates of the vector $\vec{v} = \langle 4, -7, 1\rangle$ in both bases.

   $a_1\vec{u_1} + a_2\vec{u_2} + a_3\vec{u_3} =< 4, -7, 1 >$
   $b_1\vec{z_1} + b_2\vec{z_2} + b_3\vec{z_3} =< 4, -7, 1 >$

   We augment the matrices of $\vec{u_1}, \vec{u_2}, \vec{u_3}$ and $\vec{z_1}, \vec{z_2}, \vec{z_3}$ with the vector $\vec{v}$ to find the constants, which when multiplied with respective vectors give us the coordinates of $\vec{v}$ in B_1 and B_2.

Coordinates in B_1:

$$\begin{bmatrix}1&1&0\\2&3&1\\0&2&3\end{bmatrix}\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}4\\-7\\1\end{bmatrix}$$
$$\begin{bmatrix}1&1&0&|&4\\2&3&1&|&-7\\0&2&3&|&1\end{bmatrix}=\begin{bmatrix}a_1\\a_2\\a\end{bmatrix}$$
$$\begin{bmatrix}1&1&0&|&4\\2&3&1&|&-7\\0&2&3&|&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&0&|&50\\0&1&0&|&-46\\0&0&1&|&31\end{bmatrix}$$

This means that the constants in the basis B_1 are $50, -46, 31$. Now we multiply these constants with the column vectors, to find the coordinates.

$$a_1\vec{u_1}=50*\begin{bmatrix}1\\2\\0\end{bmatrix}=\begin{bmatrix}50\\100\\0\end{bmatrix}$$
$$a_2\vec{u_2}=-46*\begin{bmatrix}1\\3\\2\end{bmatrix}=\begin{bmatrix}-46\\-138\\-92\end{bmatrix}$$
$$a_3\vec{u_3}=31*\begin{bmatrix}0\\1\\3\end{bmatrix}=\begin{bmatrix}0\\31\\93\end{bmatrix}$$

Coordinates in B_2:

$$\begin{bmatrix}1&0&1\\1&1&2\\0&1&2\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}=\begin{bmatrix}4\\-7\\1\end{bmatrix}\longrightarrow(-8,-23,12)$$
$$\begin{bmatrix}1&1&0&|&4\\2&3&1&|&-7\\0&2&3&|&1\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$$
$$\begin{bmatrix}1&1&0&|&4\\2&3&1&|&-7\\0&2&3&|&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&0&|&-8\\0&1&0&|&-23\\0&0&1&|&12\end{bmatrix}$$

The constants are $-8, -23, 12$.

$$b_1\vec{z_1}=-8*\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}-8\\-8\\0\end{bmatrix}$$

$$b_2\vec{z_2} = -23 * \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -23 \\ -23 \end{bmatrix}$$

$$b_3\vec{z_3} = 12 * \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 24 \\ 24 \end{bmatrix}$$

(c) Find the change-of-basis matrix P that sends B_1 to B_2. Check that $P[\vec{v}]_{B_1} = [\vec{v}]_{B_2}$.

We can present vector $\vec{u}$ in terms of $\vec{z}$.

$\vec{v} = 50\vec{u_1} - 46u_2 + 31u_3 = 50(c_1\vec{z_1} + c_2\vec{z_2} + c_3\vec{z_3}) - 46(e_1\vec{z_1} + e_2\vec{z_2} + e_3\vec{e_3}) + 31(f_1\vec{z_1} + f_2\vec{z_2} + f_3\vec{z_3})$

$\vec{u_1} = c_1\vec{z_1} + c_2\vec{z_2} + c_3\vec{z_3}$
$\vec{u_2} = e_1\vec{z_1} + e_2\vec{z_2} + e_3\vec{z_3}$
$\vec{u_3} = f_1\vec{z_1} + f_2\vec{z_2} + f_3\vec{z_3}$

Now we need to find the constants $c_1, c_2, c_3, e_1, e_2, e_3, f_1, f_2, f_3$.

u_1:

$$\begin{bmatrix} 1 & 0 & 1 & | & 1 \\ 1 & 1 & 2 & | & 2 \\ 0 & 1 & 2 & | & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & | & 2 \\ 0 & 1 & 0 & | & 2 \\ 0 & 0 & 1 & | & -1 \end{bmatrix}$$

$c_1 = 2, c_2 = 2, c_3 = -1$

u_2:

$$\begin{bmatrix} 1 & 0 & 1 & | & 1 \\ 1 & 1 & 2 & | & 3 \\ 0 & 1 & 2 & | & 2 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & | & 1 \\ 0 & 1 & 0 & | & 2 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$e_1 = 2, e_2 = 2, e_3 = 0$

u_3:

$$\begin{bmatrix} 1 & 0 & 1 & | & 0 \\ 1 & 1 & 2 & | & 1 \\ 0 & 1 & 2 & | & 3 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & | & -2 \\ 0 & 1 & 0 & | & -1 \\ 0 & 0 & 1 & | & 2 \end{bmatrix}$$

$f_1 = -2, f_2 = -1, f_3 = 2$

The coordinates will be multiplication of each constant with the respective vector. Now we can create a matrix of constants and prove that:

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

$\vec{v} = 50\vec{u_1} - 46u_2 + 31u_3 = 50(c_1\vec{z_1} + c_2\vec{z_2} + c_3\vec{z_3}) - 46(e_1\vec{z_1} + e_2\vec{z_2} + e_3\vec{e_3}) + 31(f_1\vec{z_1} + f_2\vec{z_2} + f_3\vec{z_3}) = \vec{v}$

$$\vec{v} = \begin{bmatrix} \\ u_1 & u_2 & u_3 \\ \\ \end{bmatrix} \begin{bmatrix} 50 \\ -46 \\ 31 \end{bmatrix} = \begin{bmatrix} \\ z_1 & z_2 & z_3 \\ \\ \end{bmatrix} \begin{bmatrix} c_1 & e_1 & f_1 \\ c_2 & e_2 & f_2 \\ c_3 & e_3 & f_3 \end{bmatrix} \begin{bmatrix} 50 \\ -46 \\ 31 \end{bmatrix} = \begin{bmatrix} \\ z_1 & z_2 & z_3 \\ \\ \end{bmatrix} *$$

$$\begin{bmatrix} 2 & 1 & -2 \\ 2 & 2 & -1 \\ -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 50 \\ -46 \\ 31 \end{bmatrix} = \begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix} \text{[Multiplied by Sage]}$$

2. Find the the change-of-basis matrix Q that sends B_2 to B_1. Check that $Q[\vec{v}]_{B_2} = [\vec{v}]_{B_1}$.

   To find the change of basis from B_2 to B_1, we take the inverse of the P, and multiply the constants by the respective vectors.

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

   $P^{-1} = Q$

   $$Q = \begin{bmatrix} 4 & -2 & 3 \\ -3 & 2 & -2 \\ 2 & -1 & 2 \end{bmatrix}$$

2 Deep Dive

1. *The least of your problems*(**#vectors**, **#linearsystems**, **#computationaltools**, **#theoreticaltools**) In this problem you will derive and then apply the least squares approximation of a solution to a system of linear equations. Suppose that you want to solve the linear system

   $$Ax = b,$$

   but the system is over-determined, i.e., it has no solution. This often happens when you have more equations than you have unknowns. Your goal is to find a vector $\hat{x}$ that is as close as possible to being a solution to $Ax = b$. The approach we take avoids the use of calculus, and is often used with large systems.

(a) Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$.

i. Why is this system inconsistent?

A. Augment the matrix by b, and turn it into $RREF$ to observe the solutions:

$$A = \begin{bmatrix} 1 & 2 & | & 4 \\ 1 & 3 & | & 5 \\ 0 & 0 & | & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \\ 0 & 0 & | & 1 \end{bmatrix}$$

From the last row we see inconsistency, thus we can declare that there is not a point in the plane where all three lines of equations intersect.

ii. Are the columns of A linearly independent? Why or why not?

After turning the matrix into $RREF$, we see that all columns are pivot columns, thus they are linearly independent. The only the linear combination of these columns that will result in $\vec{0}$ is one with the scalar of 0 (known as the trivial solution).

iii. Make a 3d plot that includes the vector b and basis vectors for the column space of A. Be sure to label the vectors.

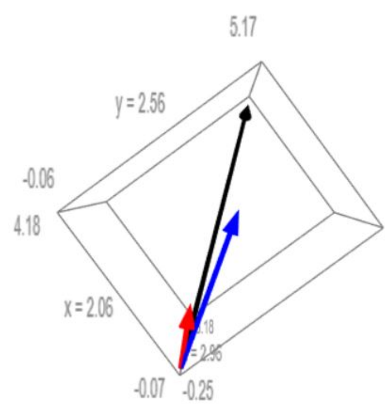

Figure 1: 3D Plotted using SAGE. Black: b, Red: a1, Blue: a2

iv. Give a geometric description of the span of the columns of A, which we will denote the "column space of A," or $Col(A)$.

$\text{span}(A) = \{a\langle 1, 1, 0\rangle + b\langle 2, 3, 0\rangle\} =$
$\{\langle a, a, 0\rangle + \langle 2b, 3b, 0\rangle\} = \{\langle a + 2b, a + 3b, 0\rangle\}$

We see that the last element of the vector that spans the space is 0, thus we can conclude that the solution lies in a plane parallel to vectors. The independence of vectors can reach any point the plane when multiplied by scalars (linear combinations).

(b) Another way to state the goal is that you want to find a vector $\hat{x}$ such that $A\hat{x}$ is as close as possible to b. In other words, you want to minimize $\|A\hat{x} - b\|$.

i. Show that $A\hat{x}$ is an element of the column space of A.

$$A\hat{x} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & 2x_2 \\ x_1 & 3x_2 \\ 0x_1 & 0x_2 \end{bmatrix} =$$

$$= \begin{bmatrix} x_1 \\ x_1 \\ 0x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 3x_2 \\ 0x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$$

As $A\hat{x}$ can be formed as a linear combinations of the vectors in A, we say that we can find it in A's column space.

A. The distance between $A\hat{x}$ and b is minimized with $b - A\hat{x}$ being orthogonal to every vector in the column space of A, i.e., $A\hat{x}$ is the projection of b onto $Col(A)$. Why?

If b has two components, one parallel and one perpendicular, then we know that we can present the parallel component from the column vector of A, or $col(A)$. What's left is the other perpendicular component which can be represent as $b - A\hat{x}$. Thus we know that $b - A\hat{x}$ is perpendicular with each component.

B. Show that this is equivalent to saying

$$A^T(b - A\hat{x}) = \vec{0}.$$

Hint: Think about matrix multiplication in terms of dot products of rows and columns.

$$A^T(b - A\hat{x}) = A^T - \frac{A^T b * A^T a}{A^T a} = \vec{0}$$

ii. We can manipulate the above formula to solve for $\hat{x}$.

A. Show that if $A^T A$ is invertible then

$$\hat{x} = (A^T A)^{-1} A^T b.$$

$$\hat{x} = (A^TA)^{-1}A^Tb$$

$$\hat{x} = A^{-1}(A^T)^{-1}A^Tb$$

...The product of inverses in reverse order

$$\hat{x} = A^{-1}b$$

...obtain identity matrix

$$A\hat{x} = AA^{-1}b$$

...Left multiply A and get identity matrix

$$b - A\hat{x} = 0$$

...re-equate by placing the unknowns in one side.

B. For the system from (a) verify that A^TA is invertible, then compute $\hat{x}$ and $A\hat{x}$.

$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix} \xrightarrow{inverse} \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$

From the equations above, we see that $\hat{x} = (A^TA)^{-1}A^Tb$, thus:

$\hat{x} = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$A\hat{x} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$

C. Add $A\hat{x}$ to your plot from (a). Discuss the result. (i.e., How does $A\hat{x}$ relate to b?)

It shows the component of b which is parallel to the column space A, hence it is a projection [Note: Plot below due to last-minute code issue].

iii. **Regression Example** Suppose you want to predict your average HC/LO score across your next day of classes based on whether you got more or less sleep than your usual 6 hours. You decide to make a regression model

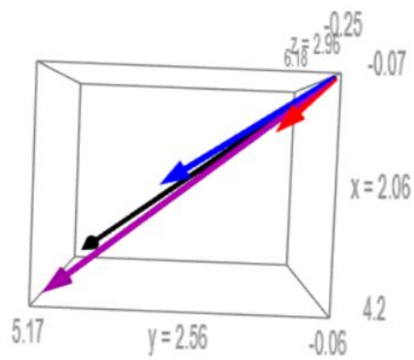

Figure 2: 3D Plotted using SAGE. Red: a1, Purple: ax, Blue: a2 and Black: b

$b = C + Mt$, where b is the output, C and M are the coefficients of your model (what you need to find), and t is the number of hours above (or below for negative values) 6 hours of sleep.
You feel lazy and only take four measurements: $b = 3$ when $t = -2$, $b = 1$ when $t = 1$, $b = 4$ when $t = 3$, and $b = 5$ when $t = 4$.

A. Inputting the above into the formula for our linear model gives us a system of four equations and two unknowns - C and M. Write out this system explicitly, then describe it as the matrix equation $Ax = b$, where $x = \begin{bmatrix} C \\ M \end{bmatrix}$. What is A?

After substituting the values from our models $b = C + Mt$, the equations are:

$3 = C - 2M$
$1 = C + M$
$4 = C + 3M$
$4 = C + 4M$

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, x = \begin{bmatrix} C \\ M \end{bmatrix}, b = \begin{bmatrix} 3 \\ 1 \\ 4 \\ 4 \end{bmatrix}$$

B. Demonstrate that the system is inconsistent and that the columns of A are independent.

Augment A with b and put it in RREF: $A = \begin{bmatrix} 1 & 2 & | & 3 \\ 1 & 1 & | & 1 \\ 1 & 3 & | & 4 \\ 1 & 4 & | & 4 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \\ 0 & 0 & | & 1 \\ 0 & 0 & | & 0 \end{bmatrix}$

We see that the system is incosistent, on the third row, and that both columns have pivots, thus they are independent.

C. Use the techniques from above to solve for $\hat{x}$.

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix} \xrightarrow{inverse} \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix}$$

$$\hat{\mathrm{x}}\text{=}(\mathrm{A}^T A)^{-1} A^T b = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$A\hat{x} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$$

D. Make a diagram that includes each of the data points and the regression line. Interpret the solution in the context of your sleep and your scores.

We can see that the data does not fit in the line perfectly, and we see that there are many outliers, but this line is the least far from each of the point. It optimizes for distance globally. For instance, if it was to optimize locally, then the line would be taking only the first data point and go through it, thus their distance would be 0. But then, the line would end up being much farther from the second, third and fourth data point. This is why this line is often called best-fit line. We can use it to make predictions. We see that this regression model is very sensitive to new data, as we do not have enough sleep information. If we add a point which is let's say much smaller than 6, which is the average right now, then the average will change drastically and we would have another line. So can really sleep really affect our scores? Probably. We need more evidence to establish a pattern, and try to draw causal inference. for right now, we can say that the more we sleep, the higher our scores. But taking other factors into account, maybe sleep is not the best predictor for our scores.

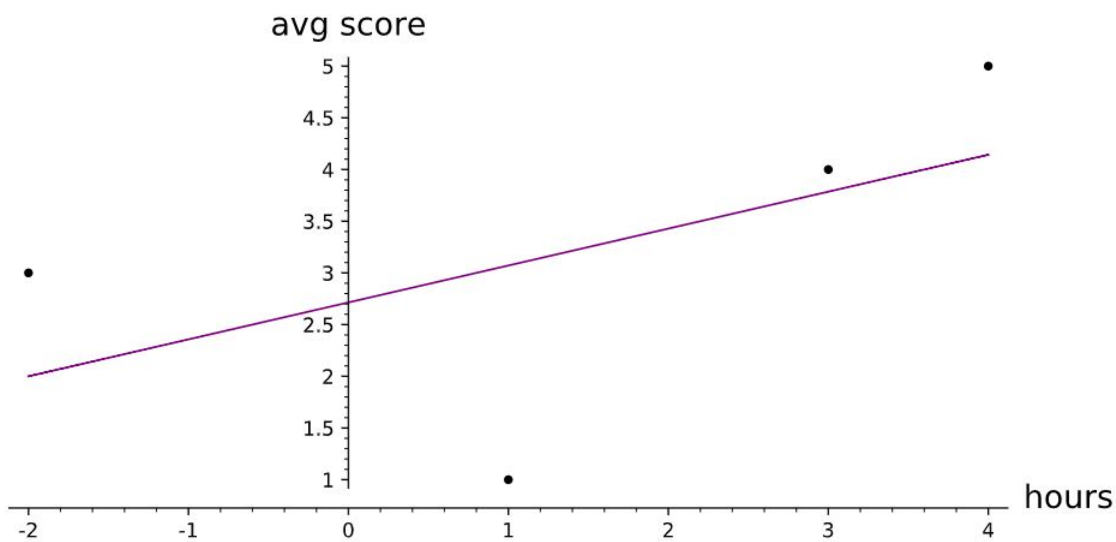

Figure 3: Plotted using SAGE. The points represent data points from the sleep data, and the line is the best fit line, in the axes of sleep hours and average score.

1. Glossary

Linear Functions - graphically these functions are straight lines. This means that all the variables are in the power of 1. As in other functions, we find values of y depending on x (or more predictors).

Linear Systems - systems of equations that contain the same variables in the power of 1 or 0. They can have a unique solution, infinitely many solutions or no solution based on whether they are consistent and parametric.

Consistent Systems - Linear Systems of Equations which contain no inconsistencies, or equations that are known mathematically not to make sense.

Augmented Matrices - represents important information such as the right-hand side of the equations or an identity matrix for the purpose of proving invertibility. We denote is as **[A|b].**

Row Operations - Transforming the matrix into another matrix of the same size requires Row Operations. We can swap rows, multiply by a scalar and add one row to the other (or the scalar of the other row).

Elementary Matrices - Elementary Matrices are matrices we use to solve Linear Systems of Equations, which are similar to identity matrices but change by a row operation that helps us reduce the original matrix to RREF.

Invertible Matrix - is a special type of matrix, which is also known as non-singular. Its determinant is different from 0. Not all matrices are invertible, but those which are, when multiplied with the original matrix should produce an identity matrix (regardless of the side of multiplication). We can find the inverse of the matrix by augmenting it with an identity matrix and reducing it to RREF.

RREF Matrix - (Reduced Row-Echelon Form) When the matrix is in this stage, we can see pivots of 1 and under and above pivots, 0s. We need those for solving linear systems when the matrix is augmented. The pivots represent the variables in a diagonal and the right-hand side represents the values of the variables for the given system. There are some criteria for a matrix to be in an RREF (more thoroughly explained, here). We can infer how many solutions the system has after turning the matrix into RREF.

Homogeneous Systems - systems where all the solutions of the system (constants or right-hand side) are 0.

Determinants - are property of matrices, more specifically a calculation of the elements of matrices based on certain rules of linearity. They are useful to make inferences about the matrices, for instance, if the determinant of the matrix is 1, we can say the system has one unique solution or if the determinant of the matrix is 0, the matrix is not invertible. We denote the determinant with **det(A),** where A represents the matrix.

Geometric Vector - a vector with length and direction in space.

Vector length: the magnitude of the vector. The magnitude of the vector $\vec{a}$ is denoted as$\|\vec{a}\|$. We find it by summing the squares of all its elements under the square root.

Vector direction - is the angle that the line makes with the horizontal segment.

Dot Product - the sum of respective components of two or more vectors (or one vector with itself). The result is a scalar. It requires two vectors of same length.

Standard Unit Vector - a vector with magnitude 1. We use them to determine direction. Each vector can be represented in terms of standard unit vectors.

Projection of Vectors - presenting one of the vectors in a straight line parallel to the other vector. We find it by the dot product of the two vectors and the unit vector defining the direction of the second vector.

Vector space - a set/space of vectors that share a set of properties such as:

1. Closed under addition

2. Closed under scalar multiplication

3. There is a zero vector $\vec{0}$ such that $\vec{0} + \vec{x} = \vec{x}$.

4. There are additive inverses.

6. The addition is commutative.

7. Addition and scalar multiplication are associative.

8. Addition and scalar multiplication have distributive properties.

Vector subspace - a vector space has a subspace when the subspace meets the three first properties of vector spaces.

Vector span - is a set of all possible combinations of the vectors in S, if S is a subspace of vector space V.

Linear Independence - is a linear combination of the vector that will result in 0, thus none of the vectors can be represented as a linear combination of others.

Basis of vectors - is a set of vectors that span a vector space and are linearly independent simultaneously.

Dimensions of vectors - the number of bases in a vector space.

Orthonormal Vectors - A set of vectors is said to be orthogonal if each of the vectors has a length of 1 and their dot product is zero, thus they form a 90-degree angle with each other.

Orthonormal Basis - is a basis for a vectors space whose vectors are orthonormal vectors.

Gram-Schmidt Algorithm - is an algorithm for orthonormalising a set of vectors.

Orthogonal Matrices - is a square matrix whose columns are orthogonal unit vectors (orthonormal vectors).

Change of Basis - representing the vectors of one basis in terms of the vectors in the other basis.

2. Appendix

#optimization: I have thoroughly explained in the last figure that the best-fit-line tries to optimize globally on distance while skipping the chance of fitting one dot into the line (optimizing locally), as that would be leading to something which is not the global optimum for the context (not showing the least distance between all dots).

#regression: I have identified a predictor (sleep hours) and the response variable (average score) and examined their relationship. I have stated that the small data we have do not give us enough information over the relationship, but from what we have we can tell that the more one sleeps, the better grades they have (positive association). I have noted the limitations of the model in the range and small dataset (sensitivity to outliers).

3. Solutions in Sage

https://cocalc.com/5f00d0c1-631f-4adc-9672-d71d8d016624/raw/A2-CS111B.sagews