Linear Transformations

A linear transformation T : V ---> W from one vector space to another is a function T which satisfies the following:

$T(v_1 + v_2) = T(v_1) + T(v_2)$ for all vectors v_1 and v_2 in V

$T(av) = aT(v)$ for all vectors v and real/complex numbers a.

(Note that as a consequence of this is that T(0) = 0 that is the zero vector in V maps to that in W)

Examples:

1. Reflection in x-axis of a point in space:
T(x, y, z) = T(x, -y, -z) is a linear transformation as it satisfies the axioms above.

2. Reflection in the origin of a point on the 2-d plane is a linear transformation
T(x, y) = (-x, -y)

3. Shear is a linear transformation
T(x, y) = (x, ax + y)

4. Rotation is a linear transformation on 2x1 matrices (in other words points in R^2) $T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

That is the if a > 0 then the point (x, y) is rotated about the origin by angle of *a* radians in the anti-clockwise direction. Otherwise it is rotated by |*a*/ radians in the clockwise direction.

5. Translation by a fixed non-zero vector v_0 is NOT a linear transformation since
$T(0) = v_0$ which violates the fundamental fact that linear transformations do NOT move the origin, that is T(0) = 0.

6. T(x, y, z) = (x + y, y + z) is a linear transformation from R^3 to R^2.

7. $T(x, y, z) = (x + y, z + 1)$ is not a linear transformation since it moves the origin

8. $T(x, y) = (x^2, y^2)$ is also not a linear transformation since it does not satisfy $T(av) = aT(v)$, for instance: $T(2(1, 1)) = (4, 4)$ while $2\,T(1, 1) = (2, 2)$.

9. Let P be the space of polynomials of degree 5 or less and Q, the same with degree 4 or less. Define T : P ---> Q as $T(f(x)) = f'(x)$ that is $f(x)$ is mapped to its derivative.

10. Given a matrix A of m x n size, the linear transformation $T : R^n \text{ ---> } R^m$ defined by $T(v) = Av$ for all vectors v in the n-dimensional space, is a linear operator.

More on Linear Transformations and Linear Operators

If V is a vector space and T : V ---> V is a linear transformation, we call T a linear operator on V.

11. The transformation T : P ---> P with T and P as above is a linear operator
12. The transformations given by rotation, reflection, shear and scaling are all linear operators on the n-dimensional space.
13. Given a **square** matrix A of n x n size, the linear transformation T : R^n ---> R^n defined by T(v) = Av for all vectors v in the n-dimensional space, is a linear operator.
14. The map T : R^n ---> R^m given by T(v) = Av + v_0 where v_0 is a fixed vector is NOT a linear transformation until and unless v_0 = 0.
15. Also, with T as above, T(v) – T(0) is a linear transformation since T(v) – T(0) = Av.
16. It can be checked that the fourier transform is a linear operator on the set of all integrable functions.

Let T : V ---> W be a linear transformation.

The **domain** of T is V.

The **co-domain** of T is W.

The **range** of T is the subspace {w in W : T(v) = w for some v in V}.

The **rank** of T is the dimension of the range of T as a subspace of W.

The **null space** of T is the subspace {v in V : T(v) = 0}.

The **nullity** of T is the dimension of the null space of T.

Transformations Represented by Matrices

Here we represent linear transformations by matrices with respect to certain given bases. We define the matrix M_T of a linear transformation T: V ---> W with respect to the ordered bases $B_1 = \{v_1, v_2, ..., v_n\}$ of V and $B_2 = \{w_1, w_2, ..., w_m\}$ of W is defined as M_T where

$$M_T = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} \text{ where } \sum_{j=1}^{m} c_{ji} w_j = T(v_i) \quad \forall 1 \le i \le m, 1 \le j \le n.$$

That is, the matrix M_T has encoded in its i^{th} column, the expression of $T(v_i)$ as a linear combination of the ordered basis vectors w_j.

Examples :

1. Find the matrix of the linear transformation T(x, y) = T(x – 2y,x + y, y – 2x) with respect to the standard bases in R^2 and R^3

   Sol. T(1, 0) = (1, 1, -2) and T(0, 1) = (-2, 1, 1) as per the definition of the linear transformation. Below the computation will be shown elaborately so that it becomes clear that there is a quick method which is also legit!

   Now we express as a linear combination as follows:

   T(1, 0) = 1 (1, 0, 0) + 1 (0, 1, 0) + (-2)(0, 0, 1)

   T(0, 1) = (-2) (1, 0, 0) + 1 (0, 1, 0) + 1 (0, 0, 1)

   Once this is done, we write the coefficients in transposed fashion as below

   $$M_T = \begin{pmatrix} 1 & -2 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}$$

   thus the above matrix is the matrix of T with respect to the standard ordered bases of R^2 and R^3. This example illustrates that the coefficients of x in each component of T(x, y) form the first column and those for y form the second column. This is not a coincidence! The matrix w.r.t the standard ordered bases is always obtainable like this for all linear transformations T: $R^n \dashrightarrow R^m$.

2. Find the matrix of T(x, y, z) = (x + y, x – y + z) w.r.t the standard ordered bases. We write out all coefficients, zero or not, as follows: T(x, y, z) = (1x + 1y + 0z, 1x – 1y + 1z)

   Thus the matrix is $M_T = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

3. If A is an m x n matrix and T : R^n ---> R^m is the linear transformation defined by
   T(v) = Av, find the matrix of T w.r.t. the standard ordered bases of R^n and R^m.

   Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$.

   Let e_i denote the vector in R^m or R^n as appropriate, where at the i^{th} position is 1 and at the rest of the positions is 0. As can be easily verified we have

   $$Ae_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{pmatrix} = \sum_{j=1}^{m} a_{ji} e_j$$

   from which it is clear by definition that the matrix of T(v) = Av w.r.t the standard ordered bases is simply A.

   NOTE: This seemingly trivial exercise is fundamental to computing the matrix of a linear transformation w.r.t any given bases.

4. Suppose B_1 and B_2 are the matrices whose columns are given by the ordered bases $\{v_1, v_2, \ldots v_n\}$ for R^n and $\{w_1, w_2, \ldots w_m\}$ for R^m respectively and T : R^n ---> R^m is given by T(v) = Av where A is a

given m x n matrix, what is the matrix of T w.r.t the ordered bases given?

The matrix M_T involves solving the systems

$$Av_i = T(v_i) = \sum_{j=1}^{m} c_{ji} w_j$$

which amounts to the matrix multiplication

$$\begin{pmatrix} \vdots & \vdots & & \vdots \\ w_1 & w_2 & \cdots & w_m \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{mi} \end{pmatrix} = \begin{pmatrix} \vdots \\ Av_i \\ \vdots \end{pmatrix} .$$

This equation above holds for the i[th] column no matter what i.

Now, including all the columns, the equation above becomes

$$\begin{pmatrix} \vdots & \vdots & & \vdots \\ Av_1 & Av_2 & \cdots & Av_n \\ \vdots & \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & & \vdots \\ w_1 & w_2 & \cdots & w_m \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}$$

The left side above is simply AB_1 and the right side is B_2M_T.

Thus we arrive at $AB_1 = B_2M_T$.

Eventually we see that $M_T = B_2^{-1}AB_1$ which is the final answer.

NOTE: To compute it, if we write down the augmented matrix $[B_2 : AB_1]$ and apply row-operations to make the left side the identity (which is as good as multiplying the left side by B_2^{-1}), then the augmented matrix becomes $[I : B_2^{-1}AB_1]$ which is nothing but $[I : M_T]$ from which we can simply read off M_T. Example 5 clarifies this.

5. E.g., Compute the matrix of T(x, y, z) = (x + y, y + z, z + x, x + y + z) w.r.t the ordered bases {(1, 0, 0), (1, 1, 0), (1, 1, 1)} and {(0, 0, 0, 1), (0, 0, 1, 1), (0, 1, 1, 1), (1, 1, 1, 1)}.

   First we find the matrix A w.r.t the standard ordered bases which is as simple as reading off coefficients like in example 2 above:

$$A=\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\\1&1&1\end{pmatrix}\ .$$

The matrix B_1 with columns the basis of the domain is

$$B_1=\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}\ .$$

The matrix B_2 with columns the basis of the range is

$$B_2=\begin{pmatrix}0&0&0&1\\0&0&1&1\\0&1&1&1\\1&1&1&1\end{pmatrix}\ .$$

To find the matrix M_T w.r.t these ordered bases we compute AB_1 and get:

$$AB_1=\begin{pmatrix}1&2&2\\0&1&2\\1&1&2\\1&2&3\end{pmatrix}\ .$$

Start with $[B_2 : AB_1]$ and row-reduce to $[I : M_T]$ as stated at the end of example 3:

$$[B_2 : AB_1]=\begin{pmatrix}0&0&0&1&:&1&2&2\\0&0&1&1&:&0&1&2\\0&1&1&1&:&1&1&2\\1&1&1&1&:&1&2&3\end{pmatrix}\ .$$

Using operations R_4 <--> R_1 and R_2 <--> R_3 we arrive at

$$\begin{pmatrix}1&1&1&1&:&1&2&3\\0&1&1&1&:&1&1&2\\0&0&1&1&:&0&1&2\\0&0&0&1&:&1&2&2\end{pmatrix}\ .$$

Using operations R_1 <--- $R_1 - R_2$, R_2 <--- $R_2 - R_3$ and R_3 <--- $R_3 - R_4$ in that order,

$$\left(\begin{array}{cccccccc} 1 & 0 & 0 & 0 & : & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & : & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & : & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & : & 1 & 2 & 2 \end{array}\right).$$

Thus we obtain the matrix

$$M_T = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ -1 & -1 & 0 \\ 1 & 2 & 2 \end{pmatrix}.$$

6. Find the matrix of the linear transformation T : P ---> P given by polynomial differentiation in the standard basis $\{1, x, x^2, x^3\}$ of the space P of polynomials of degree 3 or less.

When $f = 1$, $T(f) = 0 \quad = 0(1) + 0(x) + 0(x^2) + 0(x^3)$.
When $f = x$, $T(f) = 1 \quad = 1(1) + 0(x) + 0(x^2) + 0(x^3)$.
When $f = x^2$, $T(f) = 2x \quad = 0(1) + 2(x) + 0(x^2) + 0(x^3)$.
When $f = x^3$, $T(f) = 3x^2 \quad = 0(1) + 0(x) + 3(x^2) + 0(x^3)$.

Note that this is not like example 2 because in example 2 we had a formula for T(x,y,z) in terms of x, y and z. But here we have an expression of $T(1)$, $T(x)$, $T(x^2)$, $T(x^3)$ in terms of 1, x, x^2 and x^3. So we need to use definition of the matrix M_T of T. Thus,

$$M_T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Rotations and Reflections

A **rotation** is a transformation given by rotating every vector on the plane by a fixed angle of *a* radians.

1. Rotation by *a* radians anticlockwise is defined by $T\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\cos a & -\sin a\\ \sin a & \cos a\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$ where anti-clockwise rotation is given by $a > 0$ and clockwise by $a < 0$.
2. Rotation by an odd multiple of 180° is negation (mapping each vector to the vector in opposite direction).
3. Rotation by a multiple of 360° is the identity transformation.

A **reflection** is a transformation given by taking the mirror image of every vector across a line or plane.

1. Reflection across x-axis is given by T(x, y) = (x, -y).
2. Reflection across y-axis is given by T(x, y) = (-x, y).
3. Reflection across the origin is given by T(x, y) = (-x, -y).
4. Reflection across a line spanned by a vector v, of a vector w is given by

$$R_v(w)=\left(\frac{2vv^T}{v^Tv}-I\right)w$$

Examples:

1. Let S be a rotation by 45 degrees counterclockwise and T be reflection in the line y = 2x. Write the matrices of S and T. Is ST = TS?

Sol. $\text{matrix of } S=\begin{pmatrix}\cos 45^o & -\sin 45^o\\ \sin 45^o & \cos 45^o\end{pmatrix}=\begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$.

The line y = 2x is spanned by vector v=(1,2), and $v^Tv = [1\ 2]*[1\ 2]^T = 5$. Thus the reflection matrix is

$$\frac{2*\begin{pmatrix}1\\2\end{pmatrix}\begin{pmatrix}1 & 2\end{pmatrix}}{\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}}-I=2\frac{\begin{pmatrix}1 & 2\\2 & 4\end{pmatrix}}{5}-I=\begin{pmatrix}2/5 & 4/5\\4/5 & 8/5\end{pmatrix}-I=\begin{pmatrix}-3/5 & 4/5\\4/5 & 3/5\end{pmatrix}.$$

The matrix of ST equals $\begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}\begin{pmatrix}-3/5 & 4/5\\4/5 & 3/5\end{pmatrix}=\begin{pmatrix}\frac{-7\sqrt{2}}{10} & \frac{\sqrt{2}}{10}\\\frac{\sqrt{2}}{10} & \frac{7\sqrt{2}}{10}\end{pmatrix}$.

The matrix of TS equals $\begin{pmatrix}-3/5 & 4/5\\4/5 & 3/5\end{pmatrix}\begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}=\begin{pmatrix}\frac{-\sqrt{2}}{10} & \frac{7\sqrt{2}}{10}\\\frac{7\sqrt{2}}{10} & \frac{-\sqrt{2}}{10}\end{pmatrix}$.

Clearly, ST is not equal to TS.

2. If we take a trigonometric dig at this then we can take our v to be a unit vector i.e., v = (cos t, sin t) so that $v^Tv = 1$. on substituting for v in above we get

$$\begin{aligned}&2\frac{vv^T}{v^T v}-I\\&=2vv^T-I\\&=2\begin{pmatrix}\cos t\\\sin t\end{pmatrix}\begin{pmatrix}\cos t & \sin t\end{pmatrix}-I\\&=2\begin{pmatrix}\cos^2 t & \cos t\sin t\\\cos t\sin t & \sin^2 t\end{pmatrix}-I\\&=\begin{pmatrix}2\cos^2 t-1 & 2\cos t\sin t\\2\cos t\sin t & 2\sin^2 t-1\end{pmatrix}\\&=\begin{pmatrix}\cos 2t & \sin 2t\\\sin 2t & -\cos 2t\end{pmatrix}\end{aligned}$$

3. Find the reflection of the vector (2,3) in the line spanned by vector (1,2).

   The reflection matrix, as before for (1,2), is

$$\begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}.$$

Thus the reflection is $\begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 6/5 \\ 17/5 \end{pmatrix}$ which is (1.2, 3.4).

Inner Products, Cosines and Angle between Vectors

The inner product of v = $(v_1, v_2, ..., v_k)$ and w = $(w_1, w_2, ..., w_k)$ is defined as the sum $v_1w_1 + v_2w_2 + ... + v_kw_k$ and is denoted by v · w.

We say that two vectors v and w are orthogonal, that is $\mathrm{v} \perp \mathrm{w}$ whenever v · w = 0. in general, the angle between two vectors is given by

$$0 \le t \le \pi, \text{ where } \cos t = \frac{\mathrm{v} \cdot \mathrm{w}}{\|\mathrm{v}\|\|\mathrm{w}\|}$$

so that in the special case where t = π/2 we have that the inner product of the two vectors in question is 0.

E.g., Find the angle between the vectors (1,1,0) and (0,1,1).

Sol.

$$0 \le t \le \pi, \text{ where } \cos t = \frac{(1,1,0)\cdot(0,1,1)}{\|(1,1,0)\|\|(0,1,1)\|} = \frac{1}{2}, so\ t = \pi/3 \text{ is the angle}.$$

E.g., Find the equation of the plane of vectors that are perpendicular to the vector (3,8,7).

Sol. The vectors (x,y,z) that are perpendicular to (3,8,7) must satisfy the inner product equation (3,8,7) · (x,y,z) = (0,0,0) that is 3x+8y+7z=0.

Orthogonal Vectors and Subspaces

Given a set of vectors $S=\{v_1, v_2, ..., v_m\}$ in a vector space V, the set $S^{\perp}$ is the set of all vectors in V that are perpendicular (or orthogonal) to all the vectors $\{v_1, v_2, ..., v_m\}$.

Important Note: We have the identity for every matrix A, that:

$C(A)^{\perp} = N(A^T)$, from which it follows that

$C(A^T)^{\perp} = N(A)$.

E.g., Find the subspace of vectors orthogonal to the vectors (2,1,3) and (5,4,6).

Sol. Set $A=\begin{pmatrix} 2 & 5 \\ 1 & 4 \\ 3 & 6 \end{pmatrix}$. Thus we are looking for $C(A)^{\perp} = N(A^T)$ which is

$N\begin{pmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \end{pmatrix}$. We find RREF to be $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{pmatrix}$ in the usual way and thus

the special solution set is $\left\{\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}\right\}$. Thus the subspace of orthogonal

vectors = span <(-2,1,1)>.

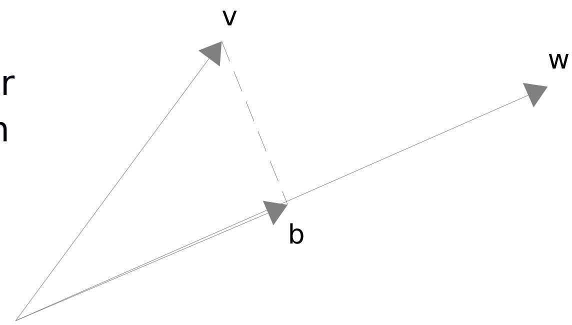
E.g., Find two independent vectors orthogonal to the intersection of the planes $2x - y + 2z = 0$ and $x - 2y + z = 0$.

Sol. Let A be the matrix $\begin{pmatrix} 2 & -1 & 2 \\ 1 & -2 & 1 \end{pmatrix}$ as given by the coefficients of the equations of the planes. The intersection of the planes is given by N(A), clearly, since N(A) consists of vectors satisfying the equations of planes.

We need vectors orthogonal to it. We know that $C(A^T)^{\perp} = N(A)$, so what we need are vectors of the row-space of A. Obvious choices of these vectors are the rows of A which are (2,-1,2) and (1,-2,1) respectively which are independent clearly.

Projection onto a Line or a Vector

The projection of a vector w onto a vector v is defined as b = cv where c minimizes ||e|| where e = w – cv, in other words, the multiple of v that is nearest in distance to w.



That is, the projection b is the vector cv so that w – cv is orthogonal to v. The vector w – b = e (error, so to speak) is denoted by the dotted line.

A precise formula for b = proj_v w is given by $b = proj_v\, w = \frac{v^T w}{v^T v} v$.

E.g., Project (1,4) onto the vector (2,3). w = (1,4). v = (2,3).

We get $b = proj_v\, w = \frac{\begin{pmatrix} 2 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 4 \end{pmatrix}}{\begin{pmatrix} 2 & 3 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix}} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \frac{14}{13}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 28/13 \\ 42/13 \end{pmatrix}$.

<u>The projection matrix</u>

The projection matrix for projection onto a vector v is given by $\frac{v v^T}{v^T v}$.

Thus $b = proj_v\, w = \frac{v v^T}{v^T v} w$.

E.g.: The projection matrix for projection onto the x-axis of the 3-d space is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ as one can quickly deduce by taking v to be the column vector (1,0,0).

E.g.: Find the projection matrix for projection onto the vector (-1,2,-2).

Sol. The projection matrix is given by

$$\frac{v\,v^T}{v^T v}=\frac{\begin{pmatrix}1\\2\\-2\end{pmatrix}\begin{pmatrix}-1 & 2 & -2\end{pmatrix}}{\begin{pmatrix}-1 & 2 & -2\end{pmatrix}\begin{pmatrix}-1\\2\\-2\end{pmatrix}}=\frac{\begin{pmatrix}-1\\2\\-2\end{pmatrix}\begin{pmatrix}-1 & 2 & -2\end{pmatrix}}{9}=\begin{pmatrix}1/9 & -2/9 & 2/9\\-2/9 & 4/9 & -4/9\\2/9 & -4/9 & 4/9\end{pmatrix}\ .$$

Projection onto a subspace

Consider a subspace spanned by a set of vectors, in other words, consider the column space of A, that is C(A). Given a vector b, we are looking for a vector v in C(A) such that $||b – v||$ is minimized.

Recall that any vector v in C(A) is of the form Ax. Thus we need to use calculus to minimize $||b – Ax||$ by taking derivatives w.r.t. x. Since the squaring is an increasing function on non-negative real numbers we may as well minimize $||b – Ax||^2$.

We know that $||b – Ax||^2 = (b – Ax)^T (b – Ax)$.

Its total derivative w.r.t. x is $-2A^T (b – Ax)$ which must be zero for critical distance (which as of now could either be minimal or maximal distance).

We equate this to zero and obtain $2A^T (b – Ax) = 0$, that is, $A^Tb = A^TAx$.

We assume that the columns of A are all independent. If not, we remove columns, leaving the independent ones that span C(A).

If the columns of A are independent then a keen linear algebra student can show that A^TA is an invertible matrix. Thus we arrive at the equation: $x = (A^TA)^{-1}A^Tb$. Since the point we are looking for is v = Ax, we have: $v = A(A^TA)^{-1}A^Tb$.

Thus the projection matrix for the projection onto C(A) is $A(A^TA)^{-1}A^T$.

Thus we have

$proj_{C(A)}(b) = A(A^TA)^{-1}A^Tb$

as our final formula.

E.g., Find the projection of the vector (1,5,4) onto the plane x+y+z=0.

Here we find vectors that span the plane x+y+z=0. We calculate the null space of the matrix [1 1 1] like before and obtain the special solutions {(-1,1,0),(-1,0,1)} which is independent since the special solutions are always an independent set.

$$\left(A^T A\right)^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix} \ .$$

Thus, $A\left(A^T A\right)^{-1} A^T = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix}$ is the projection matrix. As for the projection, we apply this to the vector (1,5,4) as follows:

$$\begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} -7/3 \\ 5/3 \\ 2/3 \end{pmatrix} \ .$$

Least Squares Line

The least squares line of best fit determines what linear model y=mx+c best fits the data points (x_i, y_i) where 1 $\leq$ i $\leq$ k with k $\geq$ 2 given.

For this, we write out equations $y_i = mx_i + c$ for each i. This in matrix form becomes:

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix} \begin{pmatrix} c \\ m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} .$$

To solve, we use $\mathrm{A} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix}, \mathrm{x} = \begin{pmatrix} c \\ m \end{pmatrix}, \mathrm{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix}$ and find the x = $\hat{\mathrm{x}}$ for which the length of Ax – b is minimized. This $\hat{\mathrm{x}} = \begin{pmatrix} c \\ m \end{pmatrix}$ gives us the coefficients in y = c + mx in the line of best fit. As in the last discussion we see that

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

E.g.: Below is a table of the height and age of a particular person taken at different stages of his life. Find the straight line that best models their relationship.

Height (inches)	Age
55	12
58	15
60	18
65	22

As per the machinery developed so far, we set $A=\begin{pmatrix}1 & 55\\ 1 & 58\\ 1 & 60\\ 1 & 65\end{pmatrix}, x=\begin{pmatrix}c\\ m\end{pmatrix}, b=\begin{pmatrix}12\\ 15\\ 18\\ 22\end{pmatrix}$

and note that the best fit coefficients are given by the vector

$$\begin{aligned}\hat{x} &= \left(A^T A\right)^{-1} A^T b\\ &= \begin{pmatrix}4 & 238\\ 238 & 14214\end{pmatrix}^{-1}\begin{pmatrix}1 & 1 & 1 & 1\\ 55 & 58 & 60 & 65\end{pmatrix}\begin{pmatrix}12\\ 15\\ 18\\ 22\end{pmatrix}\\ &= \begin{pmatrix}7107/106 & -119/106\\ -119/106 & 1/53\end{pmatrix}\begin{pmatrix}67\\ 4040\end{pmatrix}\\ &= \begin{pmatrix}-43.311\\ 1.009\end{pmatrix}.\end{aligned}$$

Thus the line of best fit is **y = 1.009x – 43.311** where y is the age and x is the height.