INF1004 L3 Linear and Affine Spaces

Alexander Binder

SIT

March 1, 2023

Outline

Linear equation system

A set of n linear equations which are to be solved for the vector $x = (x_0, \ldots x_{d-1})$

$$a_0^{\{0\}} x_0 + a_1^{\{0\}} x_1 + \ldots + a_{d-1}^{\{0\}} x_{d-1} = a^{\{0\}} \cdot x = 0 \tag{1}$$

$$a_0^{\{1\}} x_0 + a_1^{\{1\}} x_1 + \ldots + a_{d-1}^{\{1\}} x_{d-1} = a^{\{1\}} \cdot x = 0 \tag{2}$$

$$\ldots \tag{3}$$

$$a_0^{\{n-1\}} x_0 + a_1^{\{n-1\}} x_1 + \ldots + a_{d-1}^{\{n-1\}} x_{d-1} = a^{\{n-1\}} \cdot x = 0 \tag{4}$$

is called a linear equation system

Linear equation system II

A linear equation system can be defined by

$$
\begin{aligned}
Ax &= 0 \\
A.shape &= (n, d) \\
x.shape &= (d, 1)
\end{aligned}
$$

where one solves for the vector $x = (x_0, \ldots x_{d-1})$.

Affine equation systems

Affine equation system

A set of n affine equations which are to be solved for the vector $x = (x_0, \dots x_{d-1})$

$$
\begin{aligned}
a_0^{\{0\}} x_0 + a_1^{\{0\}} x_1 + \ldots + a_{d-1}^{\{0\}} x_{d-1} = a^{\{0\}} \cdot x = b_0 \\
a_0^{\{1\}} x_0 + a_1^{\{1\}} x_1 + \ldots + a_{d-1}^{\{1\}} x_{d-1} = a^{\{1\}} \cdot x = b_1 \\
\ldots \\
a_0^{\{n-1\}} x_0 + a_1^{\{n-1\}} x_1 + \ldots + a_{d-1}^{\{n-1\}} x_{d-1} = a^{\{n-1\}} \cdot x = b_{n-1}
\end{aligned}
$$

is called an affine equation system. b_i are called the bias terms.

Affine equation system II

An affine equation system can be defined by

$$\begin{aligned} Ax &= b \\ A.shape &= (n, d) \\ x.shape &= (d, 1) \\ b.shape &= (n, 1) \end{aligned}$$

where one solves for the vector $x = (x_0, \dots x_{d-1})$.

Given a matrix A of shape (n, d) and a matrix B of shape (d, f), their multiplication AB is defined as a matrix of shape (n, f) and its component $(AB)_{i,k}$ at row i and colunmn k is given as

$$(AB)_{i,k} = A_{(i,:)} \cdot B_{(:,k)} = \sum_{r=1}^{d} A_{i,r} B_{rk}$$

as an inner product between the i-th row of the left matrix and the $k-$th column of the right matrix.

Important: the number or dimensions in the second axis of A must be equal to the number or dimensions in the first axis of B. Otherwise matrix multiplication is not possible.

Therefore a matrix- matrix multiplication is a matrix consisting of inner products:

$$AB = \begin{bmatrix} A_{(0,:)} \cdot B_{(:,0)} & A_{(0,:)} \cdot B_{(:,1)} & A_{(0,:)} \cdot B_{(:,2)} & \cdots & A_{(0,:)} \cdot B_{(:,f-1)} \\ A_{(1,:)} \cdot B_{(:,0)} & A_{(1,:)} \cdot B_{(:,1)} & A_{(1,:)} \cdot B_{(:,2)} & \cdots & A_{(1,:)} \cdot B_{(:,f-1)} \\ A_{(2,:)} \cdot B_{(:,0)} & A_{(2,:)} \cdot B_{(:,1)} & A_{(2,:)} \cdot B_{(:,2)} & \cdots & A_{(2,:)} \cdot B_{(:,f-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{(n-1,:)} \cdot B_{(:,0)} & A_{(n-1,:)} \cdot B_{(:,1)} & A_{(n-1,:)} \cdot B_{(:,2)} & \cdots & A_{(n-1,:)} \cdot B_{(:,f-1)} \end{bmatrix}$$

Memorizing

- Matrix-Matrix multiplication results in a matrix, if shapes are permissible: $(n, d)(d, f) \rightarrow (n, f)$.
- The component $(AB)_{i,k}$ at row i and column k is given as the inner product between row i of the left matrix and column k of the right matrix.
  That also tells you which axes must match in dimensionality: left matrix - the number of columns $=$ the dimensionality of the second axis. right matrix - the number of rows $=$ the dimensionality of the first axis.

Matrix-Matrix multiplication

$$\begin{bmatrix} \color{blue}{1} & \color{blue}{2} & \color{blue}{3} \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} \color{blue}{5} & -1 & -2 \\ \color{blue}{3} & 9 & -3 \\ \color{blue}{-2} & 1 & 2 \end{bmatrix} = \begin{bmatrix} \color{blue}{*} & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} \color{blue}{1} & \color{blue}{2} & \color{blue}{3} \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & \color{blue}{-1} & -2 \\ 3 & \color{blue}{9} & -3 \\ -2 & \color{blue}{1} & 2 \end{bmatrix} = \begin{bmatrix} * & \color{blue}{*} & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} \color{blue}{1} & \color{blue}{2} & \color{blue}{3} \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & \color{blue}{-2} \\ 3 & 9 & \color{blue}{-3} \\ -2 & 1 & \color{blue}{2} \end{bmatrix} = \begin{bmatrix} * & * & \color{blue}{*} \\ * & * & * \\ * & * & * \end{bmatrix}$$

Matrix-Matrix multiplication

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -2 \\ 3 & 9 & -3 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -2 \\ 3 & 9 & -3 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -2 \\ 3 & 9 & -3 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -2 \\ 3 & 9 & -3 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -2 \\ 3 & 9 & -3 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & -1 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & -1 & -2 \\ 3 & 9 & -3 \\ -2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$$

Learning goals

- Understanding the solutions of linear equation systems: What is the solution which makes an inner product to zero
- Understanding the solutions of affine equation systems: What is the solution which makes an inner product to zero with an added bias term (short)
- Understanding the solutions of linear equation systems: What is the solution which makes a number of inner products to zero – alias – what happens if we intersect planes and hyperplanes
- Understanding the solutions of affine equation systems: What is the solution which makes a number of inner products with an added bias term to zero – alias – what happens if we intersect planes and hyperplanes

Outline

At the end of this we want to have an understanding how the solutions of equations such as

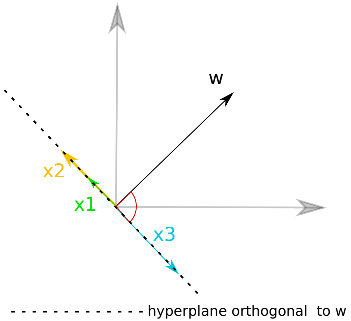
$$Ax = 0$$
$$Ax = b$$

for variable x look like. In the above A is a matrix, $x, 0, b$ are vectors.
We have seen that matrix multiplications consist of inner products.
So let us start with the question:
How do those points look like which make a single inner product become zero?

Given a fixed vector $w \in \mathbb{R}^2$ in the 2-dimensional plane, which are the the vectors x which satisfy the equation

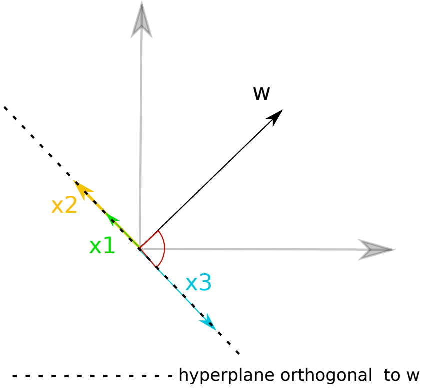
$$w \cdot x = 0$$

Obviously $x = 0$ the zero vector is a solution. What else ?



$$w \cdot x = (w_0, w_1) \cdot (x_0, x_1) = 0$$

Obviously $x = 0$ the zero vector is a solution. What else ?



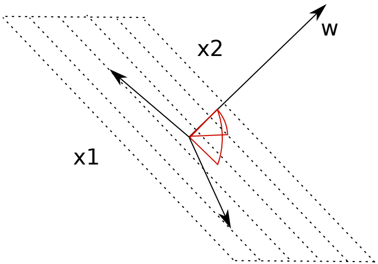
- All vectors from the origin (0 - the zero vector) to any point on the plane through the origin, which lies orthogonal to the vector w.

  $x1, x2, x3$ in the graphic are such examples.
- You can write this down explicitly in 2 dimensions:

  Given $w = (w_0, w_1)$. Then $x = (-w_1, w_0)$ is a solution, and any $ax = (-aw_1, aw_0)$, $a \in \mathbb{R}$.
- Therefore: the set of all x is a line, in case of a 2d-vector space this is a 1-dimensional vector-space.

$$w \cdot x = (w_0, w_1, w_2) \cdot (x_0, x_1, x_2) = 0$$

Obviously $x = 0$ the zero vector is a solution.
What else ?



hyperplane
through origin
orthogonal to w
vectors x1 and x2 lie inside
this hyperplane

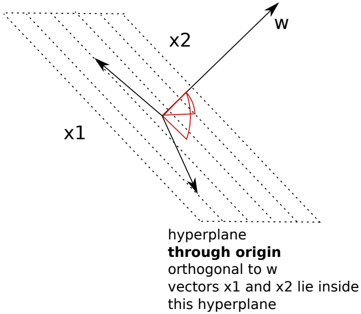
- All vectors from the origin (0 - the zero vector) to any point on the line through the origin, which lies orthogonal to the vector w.

  $x1, x2$ in the graphic are such examples.
- The set of all x is a plane, in case of a 3d-vector space, this is a 2-dimensional vector-space. Can use 2 linear independent vectors to generate all vectors of this space.
- How to obtain this for a given w?
  Use the set $v^{\{0\}} = w, v^{\{1\}} = (1, 0, 0), v^{\{2\}} = (0, 1, 0), v^{\{3\}} = (0, 0, 1)$ and run Gram-Schmidt orthogonalization on it.
- The orthogonal space (orthogonal to w) which you seek will be spanned by those two vectors from the three $v^{\{1\}}, v^{\{2\}}, v^{\{3\}}$ which did not turn into a zero vector.

Inner product = zero – case: d-dimensional vector space $\mathbb{R}^d$

$$w \cdot x = (w_0, \ldots, w_{d-1}) \cdot (x_0, \ldots, x_{d-1}) = 0$$

Obviously $x = 0$ the zero vector is a solution.
What else ?



- the analogous answer for d dimensions!

- All vectors from the origin (0 - the zero vector) to any point on the line through the origin, which lies orthogonal to the vector w.
- The set of all x is a plane. This is a $d-1$-dimensional vector-space. Can use $d-1$ linear independent vectors to generate all vectors of this space.
- How to obtain this for a given w?
  Use the set $v^{\{0\}} = w, v^{\{1\}} = e^{\{0\}}, v^{\{2\}} = e^{\{1\}}, \ldots, v^{\{d\}} = e^{\{d-1\}}$ and run Gram-Schmidt orthogonalization on it.
- The orthogonal space (orthogonal to w) which you seek will be spanned by those $d-1$ vectors from the d $v^{\{1\}}, \ldots, v^{\{d-1\}}$ which did not turn into a zero vector.

Example for $d = 5$ and $w = (1, 1, 1, 0, 0)$

$$(-1, 1, 0, 0, 0)$$
$$(0.5, 0.5, -1, 0, 0)$$
$$(0, 0, 0, 1, 0)$$
$$(0, 0, 0, 0, 1)$$

Each of the last 4 vectors is orthogonal to w and to each other. So the space of all solutions x are all linear combinations

$$x = a_0(-1, 1, 0, 0, 0) + a_1(0.5, 0.5, -1, 0, 0) + a_2(0, 0, 0, 1, 0) + a_3(0, 0, 0, 0, 1)$$

Take-away from Section 5.1

The set of vectors $x \in \mathbb{R}^d$ which solve the equation $w \cdot x = 0$ is a $d-1$-dimensional hyperplane. It can be created as linear combinations of a linearly independent set of $d-1$ vectors.

What happens if we add a bias term?

We consider now solutions x for:

$$w \cdot x = b$$

We know one solution: choose $x = b_{\|w} = b\frac{1}{w \cdot w}w$

$$x = b\frac{w}{w \cdot w}$$

$$\Rightarrow w \cdot x = b\frac{w \cdot w}{w \cdot w} = b$$

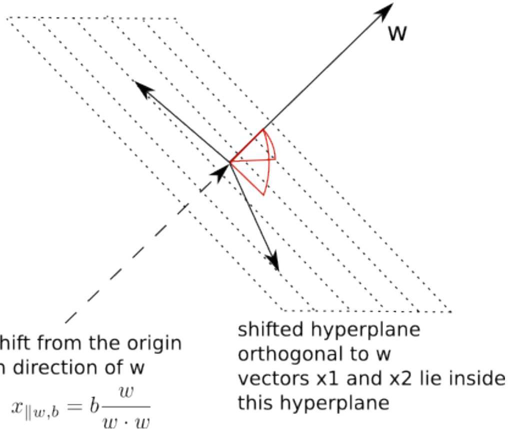
What are other solutions?

Now take any y such that $w \cdot y = 0$ and add $b_{\parallel w}$ to it.
We know then:

$$\begin{aligned} w \cdot y &= 0 \\ x &= y + b_{\parallel w} \\ \Rightarrow w \cdot x &= w \cdot (y + b_{\parallel w}) = w \cdot y + w \cdot b_{\parallel w} = 0 + b \end{aligned}$$

$$\boxed{x = y + b_{\parallel w} \text{ solves } w \cdot x = b}$$

This is the hyperplane orthogonal to w shifted into the direction of w by the vector $b_{\|w} = b\frac{1}{w \cdot w}w$.



The solution to the inner product equation with a bias term

The solution of

$$w \cdot x = b$$

is the hyperplane of all x

$$x = y + b_{\|w}, \;\; w \cdot y = 0$$

such that

- y are all vectors from the hyperplane through the origin which solve $w \cdot y = 0$,
- with the added shift vector

$$b_{\|w} = b \frac{w}{w \cdot w}$$

What is the set of all x such that $w \cdot y = b$?
In short: These x are a sum of

- a vector $b_{||w}$ which is equal to w times a constant
- and any vector y which solves $w \cdot y = 0$.

What points x make <u>a set of inner products</u> become zero ?
Extending above insights to multiple inner products

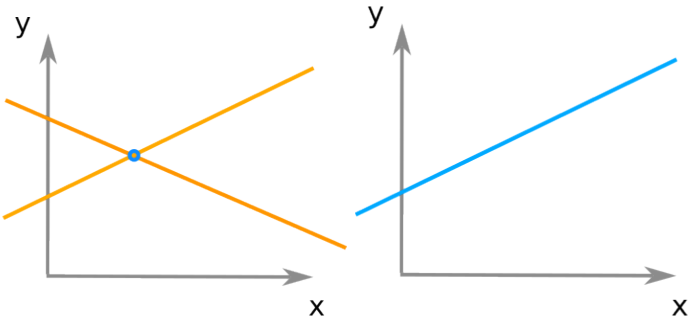
What is the intersection of two lines in $\mathbb{R}^2$ which go through the origin?

Given two fixed vectors $w^{\{0\}}, w^{\{1\}} \in \mathbb{R}^2$ in the 2-dimensional plane, which are the the vectors x which satisfy **both** equations?

$$w^{\{0\}} \cdot x = 0$$
$$w^{\{1\}} \cdot x = 0$$

We know that both equations have a line as solution.
If x satisfies both, it lies on the intersection of each solution.
This amounts to the question: what is the outcome if we intersect two lines ?

The result is:



Left: two lines do not overlap but only intersect. Right: two lines overlap perfectly.

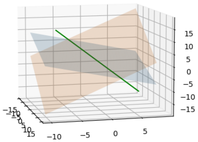
- case 2: if the lines are perfectly overlapping, then it is the line of one of the equations itself.
- case 1: otherwise: they intersect in a single point. In this case, since the zero-vector solves both equations, it must be the zero vector!

To note: the dimensionality of the solution in case 2 is reduced by 1 when intersecting the solutions.

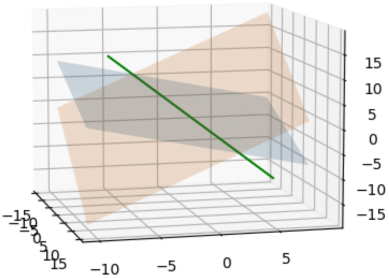
Given two fixed vectors $w^{\{0\}}, w^{\{1\}} \in \mathbb{R}^3$ in the 3-dimensional space, which are the the vectors x which satisfy both equations ?

$$x : w^{\{0\}} \cdot x = 0 \text{ and } w^{\{1\}} \cdot x = 0$$

We know that both equations have a 2-dimensional plane as solution.
So, we ask: what is the outcome if we intersect two planes ?



What is the intersection of two planes in $\mathbb{R}^3$ which go through the origin?



Two planes intersect and are not parallely aligned. The result is a line.

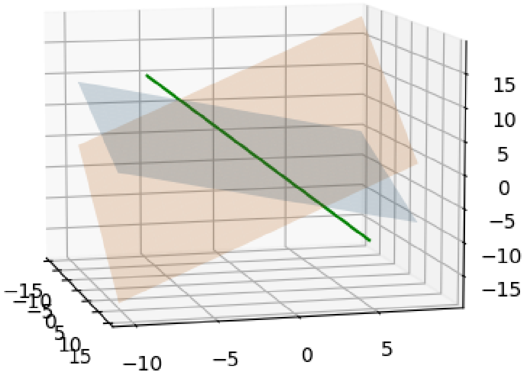
- case 1: if the two planes are perfectly parallelly aligned, then it is the 2-dim plane of one of the equations itself.
- case 2: otherwise, if they have an angle between each other – see the Figure: they intersect in a line.
  So the result would be a line, which is a 1-dimensional vector space.
- To note: the dimensionality of the solution in case 2 is again reduced by 1 when intersecting the solutions.

What is the intersection of two $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

Good news: What we have seen so far extends to higher dimensional vector spaces: if we intersect two $d-1$-dimensional planes which go through the origin, there are two cases:

- Case 1: they are perfectly parallelly aligned, then the result is the $d-1$-dimensional plane itself.
- Case 2: they have an angle between each other, then the result of the intersection of two $d-1$-dimensional hyperplanes is a $d-2$-dimensional hyperplane.

What is the intersection of two $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?



What is the intersection of two $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

How to understand that it will be $d-2$-dimensional?

- Again, use Gram-Schmidt-Orthogonalization to get there. This time apply the algorithm from the previous lesson on extending a basis to the set consisting of only the two vectors $w^{\{0\}}$ and $w^{\{1\}}$.
- start with this initialization:
  $v^{\{0\}} = w^{\{0\}}, v^{\{1\}} = w^{\{1\}}, v^{\{2\}} = e^{\{0\}}, \ldots, v^{\{3\}} = e^{\{1\}}, v^{\{d+1\}} = e^{\{d-1\}}$.
- Then we run Gram-Schmidt-Orthogonalization on this set.
- $w^{\{0\}}, w^{\{1\}}$ are not parallel, so $w^{\{1\}}$ will not become a zero vector during
- Solution: the hyperplane orthogonal to both $w^{\{0\}}$ and $w^{\{1\}}$ is spanned by those $d-2$ vectors from the set $\widetilde{v}^{\{2\}}, \ldots, \widetilde{v}^{\{d+1\}}$ which did not become zero vectors during the orthogonalization.

What is the intersection of two $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

How to understand that it will be $d-2$-dimensional?

- start with this initialization:
  $v^{\{0\}} = w^{\{0\}}, v^{\{1\}} = w^{\{1\}}, v^{\{2\}} = e^{\{0\}}, \ldots, v^{\{3\}} = e^{\{1\}}, v^{\{d+1\}} = e^{\{d-1\}}$.
- Then we run Gram-Schmidt-Orthogonalization on this set.
- $w^{\{0\}}, w^{\{1\}}$ are not parallel, so $w^{\{1\}}$ will not become a zero vector during
- Solution: the hyperplane orthogonal to both $w^{\{0\}}$ and $w^{\{1\}}$ is spanned by those $d-2$ vectors from the set $\widetilde{v}^{\{2\}}, \ldots, \widetilde{v}^{\{d+1\}}$ which did not become zero vectors during the orthogonalization.

What is the intersection of k $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

Now lets intersect $k > 2$ $d-1$-dim hyperplanes.

- We have seen before: intersecting 2 $d-1$-dim hyperplanes in d dimensions results in a hyperplane which is either $d-1$ or $d-2$-dimensional.
- How does the picture look like when intersecting k $d-1$-dim hyperplanes ? Suppose we have k vectors $w^{\{0\}}, \ldots, w^{\{k-1\}}$, and none of them is a zero vector.

What is the intersection of k $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

Now lets intersect $k>2$ $d-1$-dim hyperplanes.

The intersection of k $d-1$-dim hyperplanes in d dimensions which go through the origin

It can be a hyperplane of dimensionality anywhere between $d-1$ and $d-k$. It will be exactly $d-r$, where r is the dimensionality of the vector space spanned by $w^{\{0\}}, \ldots, w^{\{k-1\}}$.

This dimensionality r is the same as the count of the largest set of independent vectors which we can obtain from $w^{\{0\}}, \ldots, w^{\{k-1\}}$.

What is the intersection of k $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

How to see this ?

- Suppose you look at the solutions x such that

$$w^{\{0\}} \cdot x = 0,\ w^{\{1\}} \cdot x = 0, \ldots,\ w^{\{k-2\}} \cdot x = 0$$

- Suppose the set of solution vectors x for the above has dimensionality $d - r + 1$
- Suppose you add one more constraint $w^{\{k-1\}} \cdot x = 0$ to the solutions

There are two cases what can happen:

- case 1: $w^{\{k-1\}}$ is a linear combination of the previous $w^{\{0\}}, \ldots, w^{\{k-2\}}$. Then it adds no new constraint:

  Let x be a solution for the previous set:

$$w^{\{0\}} \cdot x = 0,\ w^{\{1\}} \cdot x = 0, \ldots,\ w^{\{k-2\}} \cdot x = 0$$

$$w^{\{k-1\}} = \sum_{i=0}^{k-2} a_i w^{\{i\}}$$

$$\Rightarrow w^{\{k-1\}} \cdot x = \sum_{i=0}^{k-2} a_i (w^{\{i\}} \cdot x) = 0$$

What is the intersection of k $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

How to see this ?

- Suppose you look at the solutions x such that

$$w^{\{0\}} \cdot x = 0,\ w^{\{1\}} \cdot x = 0, \ldots,\ w^{\{k-2\}} \cdot x = 0$$

- Suppose the set of solution vectors x for that is a space of dimensionality $d-r-1$
- Suppose you add one more constraint $w^{\{k-1\}} \cdot x = 0$ to the solutions

There are two cases what can happen:

- case 1: $w^{\{k-1\}}$ is a linear combination of the previous $w^{\{0\}}, \ldots, w^{\{k-2\}}$. Then it adds no new constraint:

  We will have $w^{\{k-1\}} \cdot x = 0$ for every vector x which has zero inner product for $w^{\{0\}}, \ldots, w^{\{k-2\}}$

  In this case the set of solution vectors remains to be the same space of dimensionality $d-r+1$ as before

What is the intersection of k $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

How to see this ?

- Suppose you look at the solutions x such that

$$w^{\{0\}} \cdot x = 0, \; w^{\{1\}} \cdot x = 0, \ldots, \; w^{\{k-2\}} \cdot x = 0$$

- Suppose the set of solution vectors x for that is a space of dimensionality $d - r - 1$
- Suppose you add one more constraint $w^{\{k-1\}} \cdot x = 0$ to the solutions
- case 2: $w^{\{k-1\}}$ is not a linear combination of the previous $w^{\{0\}}, \ldots, w^{\{k-1\}}$. In that case:

  It adds a new constraint to all previous solutions.

  It removes a degree of freedom from all vectors x in the direction of the newly added vector $w^{\{k-1\}}$ because we require

$$w^{\{k-1\}} \cdot x = 0$$

  Thus it reduces the dimensionality of the solution space from $d - r + 1$ to $d - r$

What is the intersection of k $d-1$-dim hyperplanes in $\mathbb{R}^d$ which go through the origin?

- Formally you can show this, if you apply the algorithm from the previous lesson on extending a set $w^{\{0\}}, \ldots, w^{\{k-1\}}$ into a basis of the vector space using Gram-Schmid
- we obtain for the first part a set of r vectors taken from $\widetilde{v}^{\{0\}}, \ldots, \widetilde{v}^{\{k-1\}}$ – because we assumed that we have r linearly independent vectors in the input $w^{\{0\}}, \ldots, w^{\{k-1\}}$
- we obtain for the second part a set of $d-r$ vectors taken from $\widetilde{v}^{\{k\}}, \ldots, \widetilde{v}^{\{k+d-1\}}$. Any linear combination from the second set provides the solution.

Outline

- One could define the rank of a set of vectors as the dimension of the space they span, which is the same as the largest number of independent vectors you can get from it . How to extend this idea to a matrix?
- You can view a matrix of shape (k, d) (that is having k rows and d columns) as a set of vectors: Either d vectors of dim k, or k vectors of dim d.

$$A = \begin{bmatrix} \begin{pmatrix} A[0,0] & A[0,1] & \dots A[0,d-1] \end{pmatrix} \\ \begin{pmatrix} A[1,0] & A[1,1] & \dots A[1,d-1] \end{pmatrix} \\ \begin{pmatrix} A[2,0] & A[2,1] & \dots A[2,d-1] \end{pmatrix} \\ \vdots \\ \begin{pmatrix} A[k,0] & A[k,1] & \dots A[k,d-1] \end{pmatrix} \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} A[0,0] \\ A[1,0] \\ \vdots \\ A[k-1,0] \end{pmatrix} & \begin{pmatrix} A[0,1] \\ A[1,1] \\ \vdots \\ A[k-1,1] \end{pmatrix} & \dots & \begin{pmatrix} A[0,d-1] \\ A[1,d-1] \\ \vdots \\ A[k-1,d-1] \end{pmatrix} \end{bmatrix}$$

The definition of matrix rank

The rank of a matrix is the largest number of independent row vectors or independent column vectors taken from the matrix.
In particular for a matrix of shape (k, d) the rank can be at most $\min(k, d)$

Matrix rank and vector space dimensionality

The matrix rank is equal to the dimensionality of the vector space spanned by rows <u>and</u> by columns of a matrix.

Outline

The solution to the linear equation set I

$$Ax = 0$$

Can be interpreted as the intersection of hyperplanes through the origin ($=$ containing the zero-vector), where

- the i-th hyperplane is orthogonal to the i-th row of A, $A[i,:]$

The solution to the linear equation set II

Let A be a matrix of shape (k, d), x a d-dim vector. The the solution to the equation

$$Ax = 0$$

is a vector space of dimensionality $d - r$ where r is the matrix rank.

The vector space of solutions x are all vectors which are orthogonal to all the row vectors of A (and r is the dimensionality spanned by the set of row vectors of A).

If $k \geq d$ and the matrix rank is d, then the solution is only $x = 0$.

What is the solution to the affine equation $Ax = b$?

Now we look at intersections of hyperplanes,

$$A[0,:] \cdot x = b_0$$
$$A[1,:] \cdot x = b_1$$
$$A[2,:] \cdot x = b_2$$

which are shifted by bias terms b_k.
What is different to the case $Ax = 0$?
The case $Ax = b$ may have no solutions.
Lets explain this on the next slides!

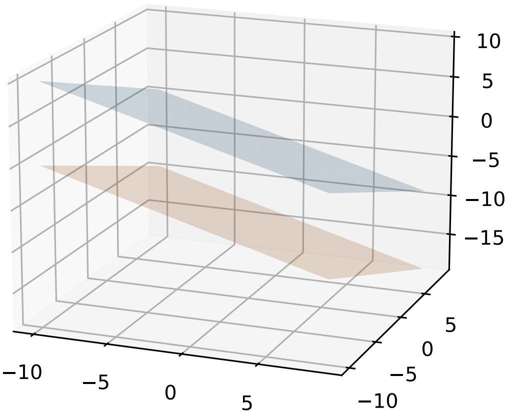
What is the solution to the affine equation $Ax = b$? When there is no solution!

A simple explanation: Lets consider the case that matrix A in $Ax = b$ has two identical rows $w = A[i, :] = A[k, :]$ but different bias values $b_i \neq b_k$. This corresponds to trying to solve

$$w \cdot x = b_i$$
$$w \cdot x = b_k$$

If $b_i \neq b_k$, then we have two parallel hyperplanes, shifted by two different amounts b_i, b_k. They intersect nowhere. So, no solution can be found.

What is the solution to the affine equation $Ax = b$? When there is no solution!



What is the solution to the affine equation $Ax = b$? When there is no solution!

A similar thought holds if two rows are not identical, but a row $A_{k,:}$ is a linear combination of other rows $\sum_{i:i\neq k} a_i A_{i,:}$.
We start with the equation for the one row which is a linear combination of other rows.

$$A_{k,:} \cdot x = b_k$$

It is a linear combination of others, therefore:

$$\begin{aligned} A_{k,:} &= \sum_{i:i\neq k} a_i A_{i,:} \\ \text{plug this in: } \sum_{i:i\neq k} a_i A_{i,:} \cdot x &= b_k \\ \text{but also: } \sum_{i:i\neq k} a_i A_{i,:} \cdot x &= \sum_{i:i\neq k} a_i b_i \end{aligned}$$

The last one from summing the other rows according to the coefficients a_i of the linear combination.

A similar thought holds if two rows are not identical, but a row $A_{k,:}$ is a linear combination of other rows $\sum_{i:i\neq k} a_i A_{i,:}$.
We obtain two conditions:

$$\sum_{i:i\neq k} a_i A_{i,:} \cdot x = b_k$$
$$\sum_{i:i\neq k} a_i A_{i,:} \cdot x = \sum_{i:i\neq k} a_i b_i$$

If one replaces $w = \sum_{i:i\neq k} a_i A_{i,:}$, then we are at the case above

$$w \cdot x = b_k$$
$$w \cdot x = \sum_{i:i\neq k} a_i b_i$$

What is the solution to the affine equation $Ax = b$? When there is no solution!

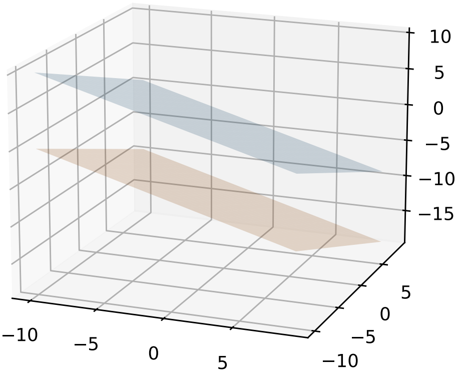
When a row $A_{k,:}$ is a linear combination of other rows $\sum_{i:i\neq k} a_i A_{i,:}$, then we obtain again a case

$$w \cdot x = b_k$$
$$w \cdot x = \sum_{i:i\neq k} a_i b_i$$

- If $\sum_{i:i\neq k} a_i b_i \neq b_k$, then this has no solution. Geometrically, in this case we have two hyperplanes which are parallel but shifted and therefore have no intersection to each other
- If $\sum_{i:i\neq k} a_i b_i = b_k$, then this may have a solution (if all other rows are consistent, too, by checking at their bias terms).

What is the solution to the affine equation $Ax = b$? When there is no solution!

We can have the case that the equations $A[i,:] \cdot x = b_i$ define hyperplanes which are parallel and have no intersection. This can happen, whenever a row $A[i,:]$ is a linear combination of other rows.



What is the solution to the affine equation $Ax = b$?

The solution to the affine equation I

$$Ax = b$$

Can be interpreted as the intersection of shifted hyperplanes, where

- the i-th hyperplane is orthogonal to the i-th row of A, $A[i, :]$, and,
- its shift depends on the i-th bias term b_i

In case that two hyperplanes are parallel but have a different shift value, no solution exist.

The solution to the affine equation II

Let A be a matrix of shape (k, d), x a d-dim vector. The the solution to the equation

$$Ax = b$$

- either does not exist
- or it is an affine space of dimensionality $d - r$ where r is the matrix rank.

An affine space is a vector space V with an added offset vector u.

- The affine space of solutions x are all vectors summed from two components $x = v + u$, such that
- v is orthogonal to all the row vectors of A, that is, the vector v solves $Av = 0$, and
- u is any vector which is solving $Au = b$.

r is the dimensionality spanned by the set of row vectors of A.

One can simply check that this is true:

$$
\begin{aligned}
Au &= b \\
Av &= 0 \\
\Rightarrow A(u+v) &= Au + Av = b + 0 = b
\end{aligned}
$$

So every $u : Au = b$, $v : Av = 0$ realize a solution by their sum $x = u + v$

- We have shown that solutions of the equation $Ax = 0$ are
  - an intersection of hyperplanes through the origin
  - which is a vector space
- We have shown that solutions of the equation $Ax = b$, if it has a solution,
  - an intersection of hyperplanes which are shifted
  - We will formalize the idea of hyperplanes which are shifted as a vector space with an added shift vector, which is called an **affine space**
- Next: we will show that we can represent any vector space as a solution of the equation $Ax = 0$.
- Also Next: we will show that we can represent any affine space as a solution of the equation $Ax = b$.

vector space

Let $v_0, \ldots, v_{d-1}$ be some vectors, t_i are real numbers.
Then the set of all linear combinations

$$f(t_0, \ldots, t_{d-1}) = \sum_{i=0}^{d-1} t_i v_i$$

is a vector space.
The dimensionality is the rank of the set $v_0, \ldots, v_{d-1}$.

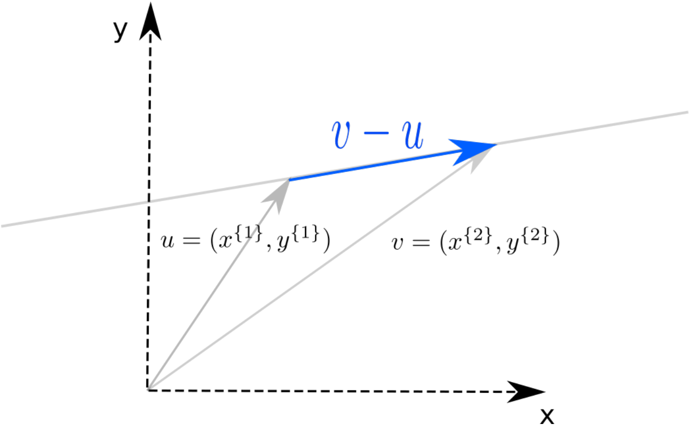
Affine space

Let $v_0, \ldots, v_{d-1}$ be some vectors, u another vector, and t_i are real numbers.
Then the set of all linear combinations with added vector u

$$f(t_0, \ldots, t_{d-1}) = u + \sum_{i=0}^{d-1} t_i v_i$$

is an affine space.
The dimensionality is the rank of the set $v_0, \ldots, v_{d-1}$.

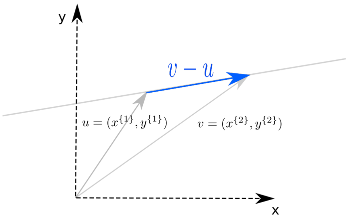
This is the special case of an 1D-affine or linear space.



If u, uv are any two points on the line, then one can represent the line as:

$$f(t) = u + t(v - u)$$
$$f(t) = u + t(u - uv)$$

Reason: $\pm(u - v)$ is the direction of the line



Representing a 1-d line

A line through the origin

$$f(t) = tz$$

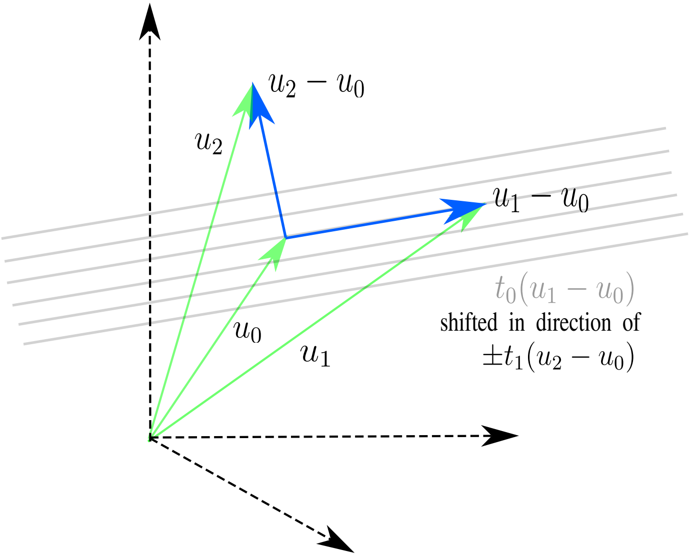
where z is the vector in direction of the line. This is a vector space.

A line passing not through the origin

$$f(t) = u + tz$$

where z is the vector in direction of the line. This is an affine space.
u is any vector on the line.

This is the special case of an 2D-affine or linear space.

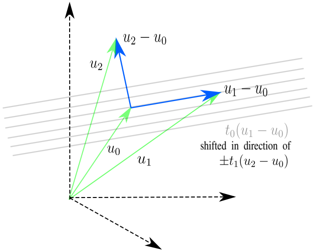


Representing a 2-d plane

A 2d plane

If u_0, u_1, u_2 are any three points on the plane such that $u_1 - u_0$ and $u_2 - u_0$ are linearly independent,
then one can represent the 2D plane as:

$$f(t_0, t_1) = u_0 + t_0(u_1 - u_0) + t_1(u_2 - u_0)$$



$u_1 - u_0$ and $u_2 - u_0$ are linearly independent, if

- there exist no $c \in \mathbb{R}$ such that $(u_1 - u_0) = c(u_2 - u_0)$
- or $(u_1 - u_0) \cdot (u_2 - u_0) \neq \pm \|u_1 - u_0\| \|u_2 - u_0\|$
- or $|(u_1 - u_0) \cdot (u_2 - u_0)| < \|u_1 - u_0\| \|u_2 - u_0\|$
- or $\cos \angle(u_1 - u_0, u_2 - u_0) \neq \pm 1$,

Outline

Represent a vector or affine space as the solution of a linear or affine equation system

We go now beyond lines and 2d planes. We seek to find a matrix A and a vector b such that each element $f(t_0, \dots t_{d-1})$ of the affine space solves $Af(t_0, \dots t_{d-1}) = b$

We start with

$$f(t_0, \dots t_{d-1}) = u + \sum_{i=0}^{d-1} t_i v_i$$

- step 1: Let k be the rank of the set $\{v_0, \dots, v_{d-1}\}$, e is the dimension of the vector space. Then find $e - k$ linearly independent vectors w_j forming a set $\{w_0, \dots, w_{e-k-1}\}$ such that $w_j \cdot v_i = 0$.

Next: How to achieve step 1?

Represent a vector or affine space as the solution of a linear or affine equation system: step 1

Obtain ($e - k$ many) linear independent solutions $\{w_0, \ldots, w_{e-k-1}\}$ to the linear equation system (see solving linear equation systems, not plugging in 0 for $*$)

$$B = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{d-1} \end{bmatrix}$$

$$Bw = 0$$

Why ? Every solution w for $Bw = 0$ obviously satisfies $w \cdot v_i = 0$ by definition of matrix multiplication:

$$Bw = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{d-1} \end{bmatrix} w = \begin{bmatrix} v_0 \cdot w \\ v_1 \cdot w \\ \vdots \\ v_{d-1} \cdot w \end{bmatrix} = 0$$

Represent a vector or affine space as the solution of a linear or affine equation system: step 2

- step 2: define A using the solutions from step 1

$$A = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{e-d-1} \end{bmatrix}$$

and set

$$b = Au$$

then all vectors in $f(t_0, \dots t_{d-1})$ are the solution of the affine equation system

$$Ax = b$$

Why does this work?

$$
\begin{aligned}
Af(t_0, \dots t_{d-1}) &= A(u + \sum_{i=0}^{d-1} t_i v_i) = Au + A(\sum_{i=0}^{d-1} t_i v_i) \\
&= b + \sum_{i=0}^{d-1} t_i A v_i \\
&= b + \sum_{i=0}^{d-1} t_i \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{e-d-1} \end{bmatrix} v_i \\
&= b + \sum_{i=0}^{d-1} t_i \begin{bmatrix} w_0 \cdot v_i \\ w_1 \cdot v_i \\ \vdots \\ w_{e-d-1} \cdot v_i \end{bmatrix}
\end{aligned}
$$

Why does this work?

$$Af(t_0, \ldots t_{d-1}) = A(u + \sum_{i=0}^{d-1} t_i v_i) = Au + A(\sum_{i=0}^{d-1} t_i v_i)$$

$$= b + \sum_{i=0}^{d-1} t_i \begin{bmatrix} w_0 \cdot v_i \\ w_1 \cdot v_i \\ \vdots \\ w_{e-d-1} \cdot v_i \end{bmatrix}$$

$$= b + \sum_{i=0}^{d-1} t_i \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = b$$

Outline

Vector space recap

Remember: Vector space as an abstract concept a set of elements v

- which we can multiply with a real value
- which we can add together
- such that multiplications and additions stay within the set
- it needs also a neutral element $0 + v = v$

We know this works for vectors of real numbers $v = (a_0, a_1, a_2)$.

We have shown above I

- The set of solutions x for $Ax = 0$ is a vector space, that is, one can express it as: $x = \sum_{i=0}^{d-1} a_i v_i$ for some set of basis vectors $\{v_0, \ldots, v_{d-1}\}$ and all possible real numbers a_i
- A parametrization $x = \sum_{i=0}^{d-1} a_i v_i$ can be converted into $Ax = 0$ by finding a suitable matrix A

We have shown above II

- The set of solutions x for $Ax = b$ is an affine space (if it has solutions), that is, one can express it as: $x = u + \sum_{i=0}^{d-1} a_i v_i$ for some set of basis vectors $\{v_0, \ldots, v_{d-1}\}$, all possible real numbers a_i, and a translation vector u
- A parametrization $x = u + \sum_{i=0}^{d-1} a_i v_i$ can be converted into $Ax = b$ by finding a suitable matrix A and an offset vector b

There is a correspondence:

set	spanned by	parametrization	Equation
Vector sp	$\{v_0, \ldots, v_{d-1}\}$	all $x = \sum_{i=0}^{d-1} a_i v_i$ for all possible real a_i	$Ax = 0$
Affine sp	u ,$\{v_0, \ldots, v_{d-1}\}$	all $x = u + \sum_{i=0}^{d-1} a_i v_i$ for all possible real a_i	$Ax = b$

Does this fit to the cases without or with only one solution?

The vector space with only the zero vector:

$$V = \{0\}$$

is a vector space. $a0 + b0 = 0$ for all elements from this space :)
It is zero-dimensional.
Therefore $V = \{u\} = u + \{0\}$ is an affine space.

What about $Ax = b$ without solutions?
The empty set is not a vector space. It has no 0 vector.

Outline

Definition of a Linear mapping

A function $f : V \to Y$ which takes a vector $v \in V$ and outputs a vector $f(v) \in Y$ is a linear mapping if

$$f(cv) = cf(v)$$
$$f(v + w) = f(v) + f(w)$$

holds

Intuition of a Linear mapping

- you can swap linear mappings with multiplication with a constant

$$f(cv) = cf(v)$$

  Linear mapping of a multiplication = multiplication of the linear mapping

- you can swap linear mappings with addition of two inputs

$$f(v + w) = f(v) + f(w)$$

  Linear mapping of a sum = sum of linear mappings

Inner products can be used to define linear mappings:

$$f_z(v) = z \cdot v$$

is a linear mapping in the vector argument v.
Reason:

$$f_z(cv) = z \cdot (cv) = \sum_i z_i c v_i = cz \cdot v = cf_z(v)$$

$$f_z(v_0 + v_1) = z \cdot (v_0 + v_1) = \sum_i z_i (v_{0,i} + v_{1,i}) = \sum_i z_i v_{0,i} + \sum_i z_i v_{1,i} = f_z(v_0) + f_z(v_1)$$

Matrix-vector multiplications define linear mappings (two ways!)

Way 1:

$$g_A(v) = Av$$

is a linear mapping in the vector argument v.
Way 2:

$$h_v(A) = Av$$

is a linear mapping in the matrix argument A.
Reason:

$$g_A(v_0 + v_1) = A(v_0 + v_1) = Av_0 + Av_1 = g_A(v_0) + g_A(v_1)$$
$$h_v(A_0 + A_1) = (A_0 + A_1)v = A_0 v + A_1 v = h_v(A_0) + h_v(A_1)$$

the same also works when we check for constants $g_a(cv) = \ldots, h_v(cA) = \ldots$.

Matrix-Matrix multiplications define linear mappings in two ways.

$$d_A(B) = AB$$

is a linear mapping in the matrix argument B.

$$e_B(A) = AB$$

is a linear mapping in the matrix argument A.

Matrix-Matrix and Matrix-vector multiplications define linear mappings

$g_A(v) = Av$	linear in argument v
$h_v(A) = Av$	linear in argument A
$d_A(B) = AB$	linear in argument B
$e_B(A) = AB$	linear in argument A