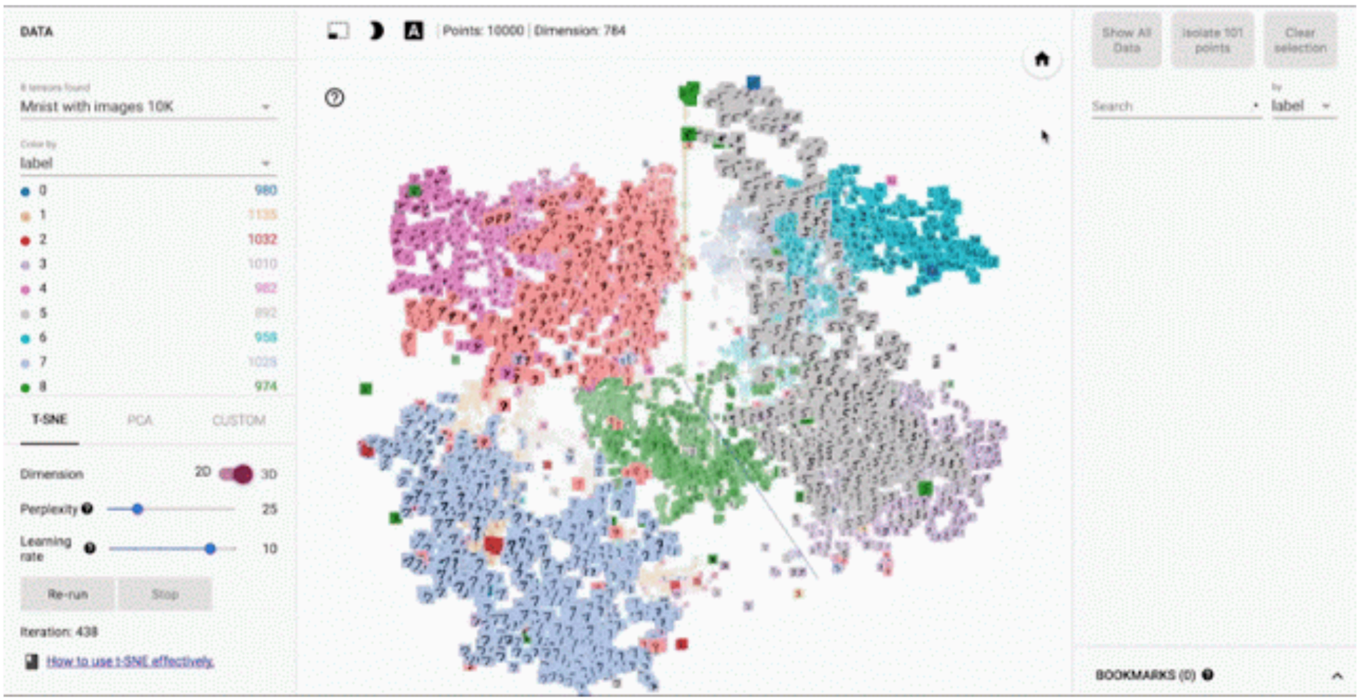
Introduction to Neural Computation

Prof. Michale Fee
MIT BCS 9.40 — 2017

Lecture 16
Networks, Matrices and Basis Sets

Seeing in high dimensions



https://research.googleblog.com/2016/12/open-sourcing-embedding-projector-tool.html

Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis

Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis

Two-layer feed-forward network

- We can expand our set of output neurons to make a more general network…

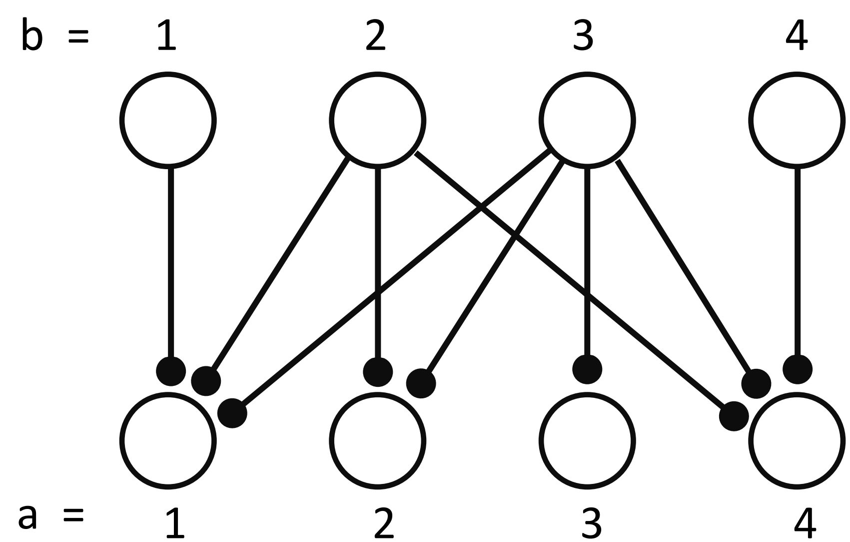
input firing rates

$$\left[u_1 , u_2 , u_3 ,\ldots\ u_{n_b} \right] = \vec{u}$$

Lots of synaptic weights! W_{ab}

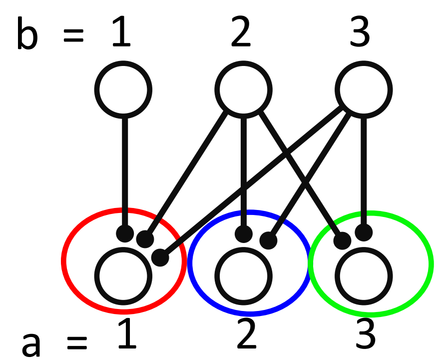
output firing rates

$$\left[v_1 , v_2 , v_3 ,\ldots\ v_{n_a} \right] = \vec{v}$$



Two-layer feed-forward network

- We now have a weight from each of our input neurons onto each of our output neurons!



- We write the weights as a matrix.

weight matrix

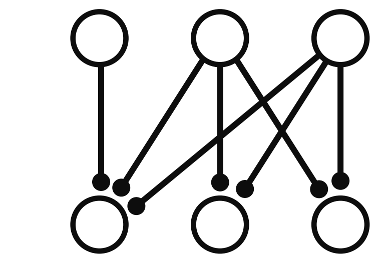
$$W_{ab} = \begin{array}{cc} & \begin{array}{ccc} b = 1 & 2 & 3 \end{array} \\ \begin{array}{c} a = 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \end{array} = \begin{bmatrix} \vec{w}_{a=1} \\ \vec{w}_{a=2} \\ \vec{w}_{a=3} \end{bmatrix}$$

a	b
row	column
post	pre

Two-layer feed-forward network

- We can write down the firing rates of our output neurons as a matrix multiplication.

$$\vec{v} = W\,\vec{u} \qquad v_a = \sum_b W_{ab} u_b$$



$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \vec{w}_{a=1} \cdot \vec{u} \\ \vec{w}_{a=2} \cdot \vec{u} \\ \vec{w}_{a=3} \cdot \vec{u} \end{bmatrix}$$

- Dot product interpretation of matrix multiplication

Two-layer feed-forward network

- There is another way to think about what the weight matrix means…

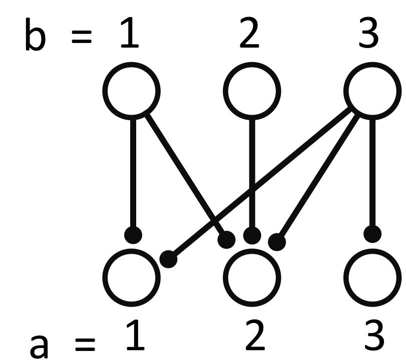
$$\vec{v} = W\vec{u} = \begin{array}{c} \begin{array}{ccc} b = 1 & \quad 2 & \quad 3 \end{array} \\ \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \end{array} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$\left[\vec{w}^{(1)} \;\middle|\; \vec{w}^{(2)} \;\middle|\; \vec{w}^{(3)} \right]$$

vector of weights from input neuron 1

vector of weights from input neuron 2

vector of weights from input neuron 3



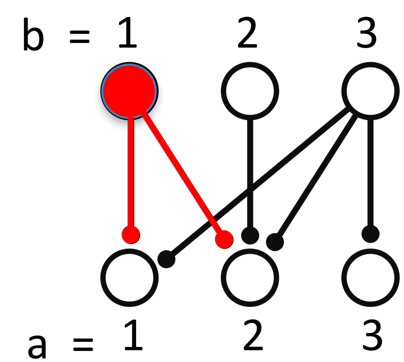
$$W = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Two-layer feed-forward network

- There is another way to think about what the weight matrix means…

$$
\vec{v} = W\vec{u} = \begin{array}{cccc} b = 1 & 2 & 3 & \\ \left[\begin{matrix} w_{11} \\ w_{21} \\ w_{31} \end{matrix}\right. & \begin{matrix} w_{12} \\ w_{22} \\ w_{32} \end{matrix} & \left.\begin{matrix} w_{13} \\ w_{23} \\ w_{33} \end{matrix}\right] & \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \end{array}
$$

$$
\left[\, \vec{w}^{(1)} \,\middle|\, \vec{w}^{(2)} \,\middle|\, \vec{w}^{(3)} \,\right]
$$



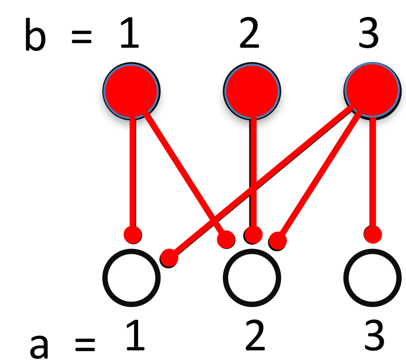
- What is the output if only input neuron 1 is active?

$$
\vec{v} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \\ 0 \end{bmatrix} = u_1 \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} = u_1\,\vec{w}^{(1)} = u_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}
$$

Two-layer feed-forward network

$$\vec{v} = W\vec{u} = \begin{array}{c} \begin{array}{ccc} b=1 & 2 & 3 \end{array} \\ \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \end{bmatrix} \end{array} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$\left[\vec{w}^{(1)} \middle| \vec{w}^{(2)} \middle| \vec{w}^{(3)} \right]$$



$$\vec{v} = u_1 \begin{bmatrix} w_{11} \\ w_{21} \\ w_{31} \end{bmatrix} + u_2 \begin{bmatrix} w_{12} \\ w_{22} \\ w_{32} \end{bmatrix} + u_3 \begin{bmatrix} w_{13} \\ w_{23} \\ w_{33} \end{bmatrix}$$

$$W = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

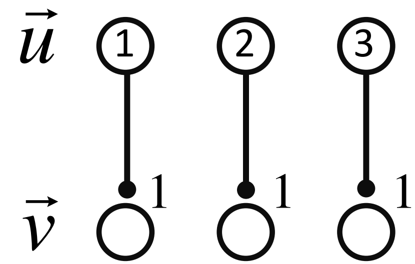
$$\vec{v} = u_1 \vec{w}^{(1)} + u_2 \vec{w}^{(2)} + u_3 \vec{w}^{(3)}$$

The output pattern is a linear combination of contributions from each of the input neurons!

Examples of simple networks

- Each input neuron connects to one neuron in the output layer, with a weight of one.

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad W = I$$



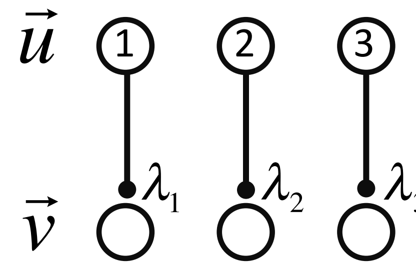
$$\vec{v} = W\vec{u} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$\vec{v} = \vec{u}$$

Examples of simple networks

- Each input neuron connects to one neuron in the output layer, with an arbitrary weight

$$W = \Lambda \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

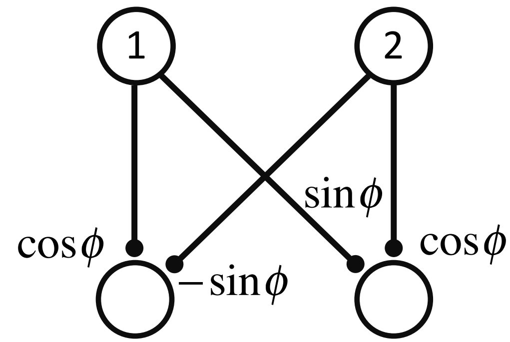


$$\vec{v} = \Lambda \vec{u} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 u_1 \\ \lambda_2 u_2 \\ \lambda_3 u_3 \end{bmatrix}$$

Examples of simple networks

- Input neurons connect to output neurons with a weight matrix that corresponds to a rotation matrix.

$$W = \mathbf{\Phi} \qquad \mathbf{\Phi} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

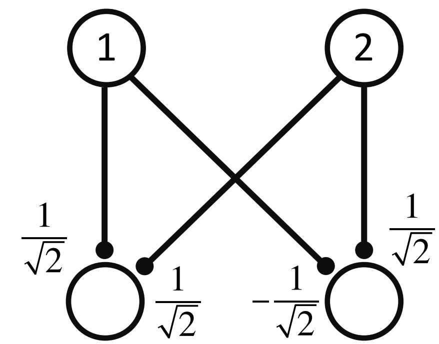


$$\vec{v} = \mathbf{\Phi} \cdot \vec{u} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \cos\phi - u_2 \sin\phi \\ u_1 \sin\phi + u_2 \cos\phi \end{bmatrix}$$

Examples of simple networks

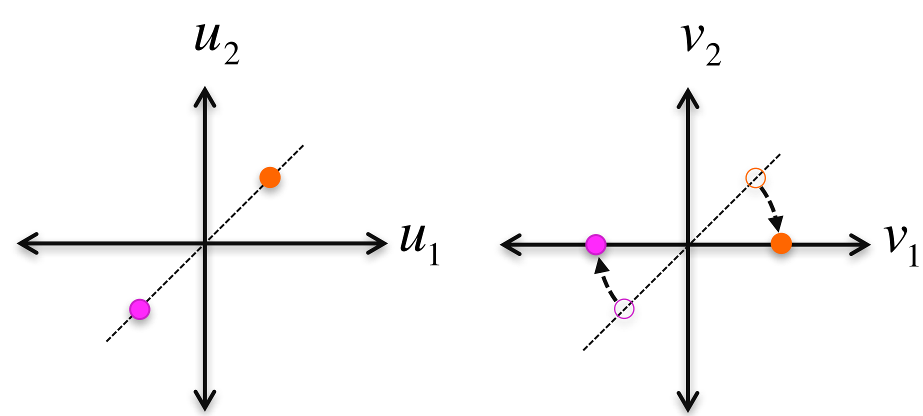
- Let's look at an example rotation matrix (ϕ=-45°)

$$\Phi(-45^0) = \begin{bmatrix} \cos(-\frac{\pi}{4}) & -\sin(-\frac{\pi}{4}) \\ \sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$



$$\vec{v} = \Phi \cdot \vec{u} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

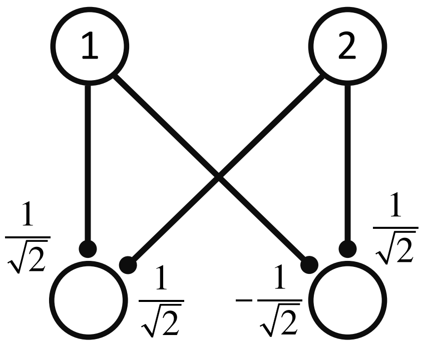
$$\vec{v} = \frac{1}{\sqrt{2}} \begin{bmatrix} u_2 + u_1 \\ u_2 - u_1 \end{bmatrix}$$

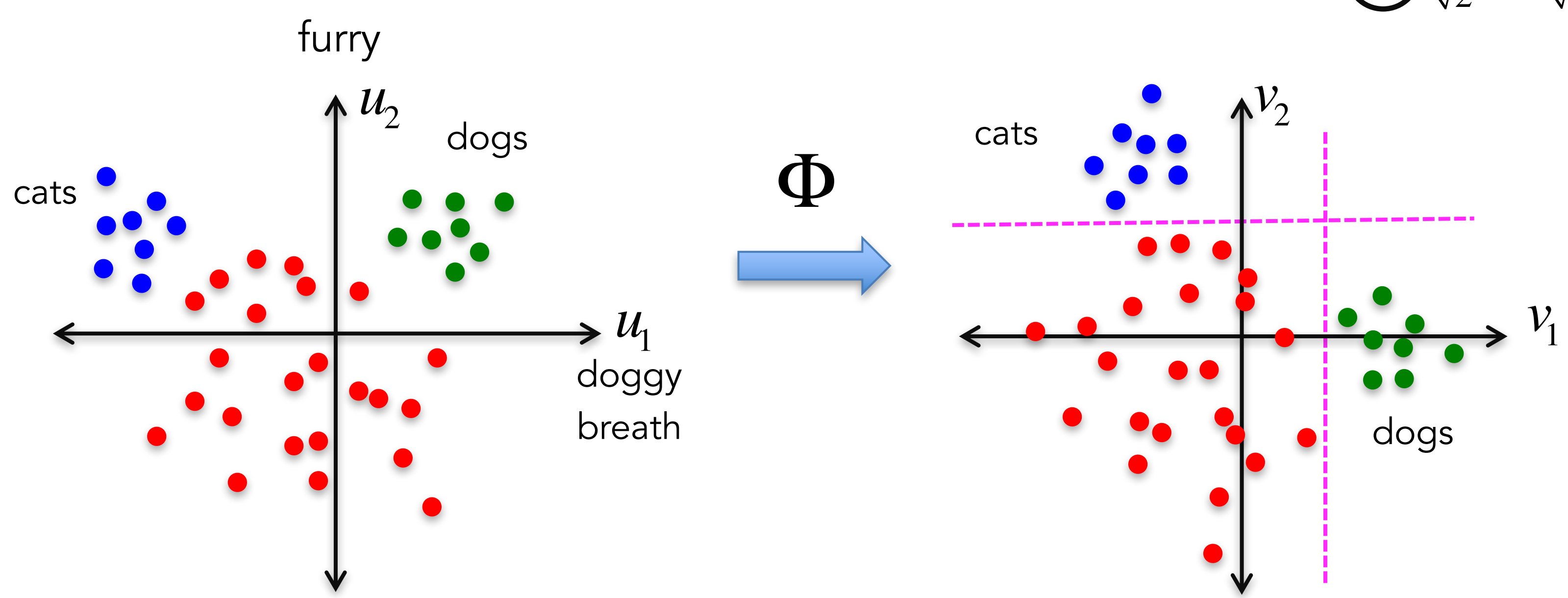


Examples of simple networks

- Rotation matrices can be very useful when different directions in feature space carry different useful information

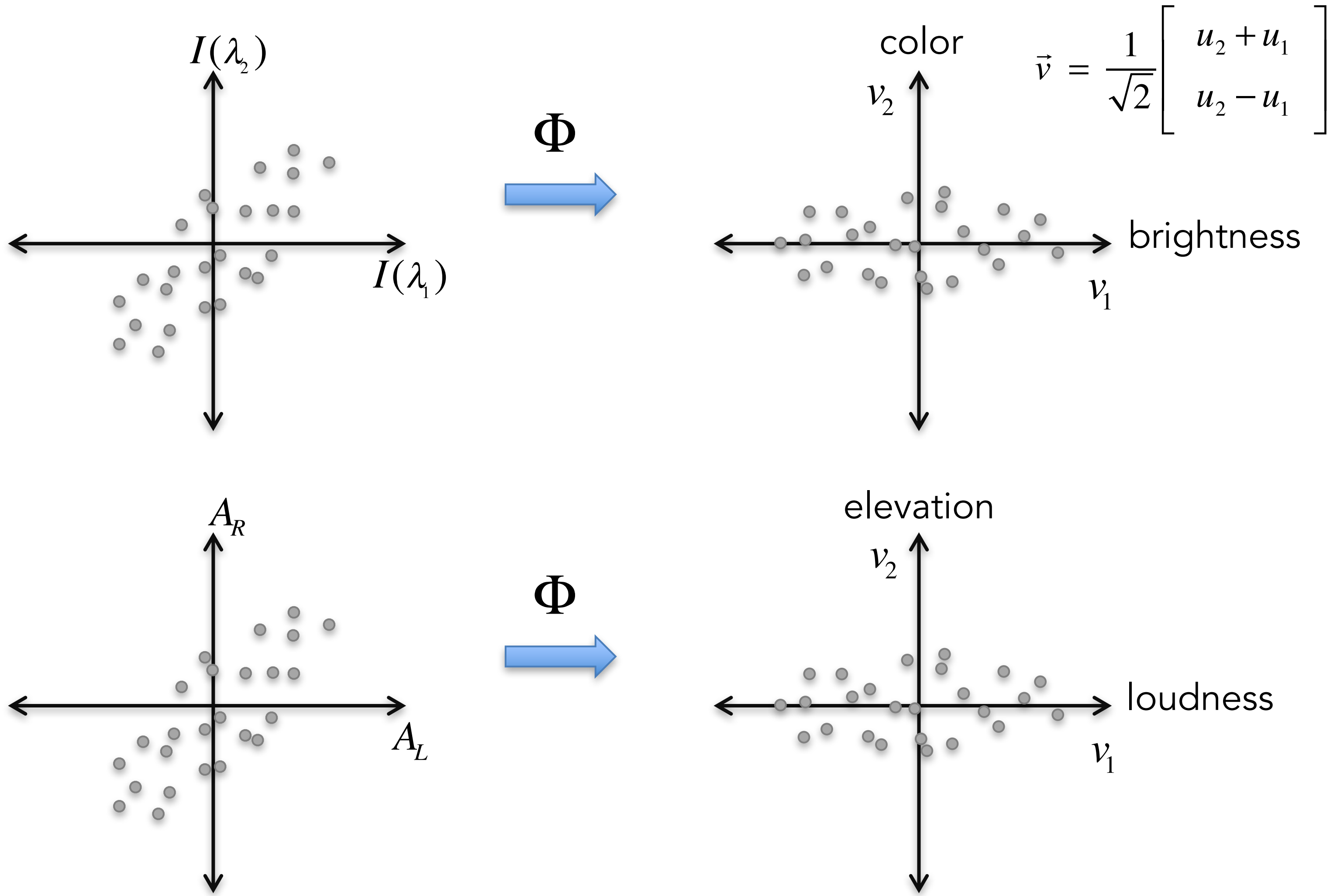
$$\Phi(-45^0) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$





Examples of simple networks

- Rotation matrices can be very useful when different directions in feature space carry different useful information



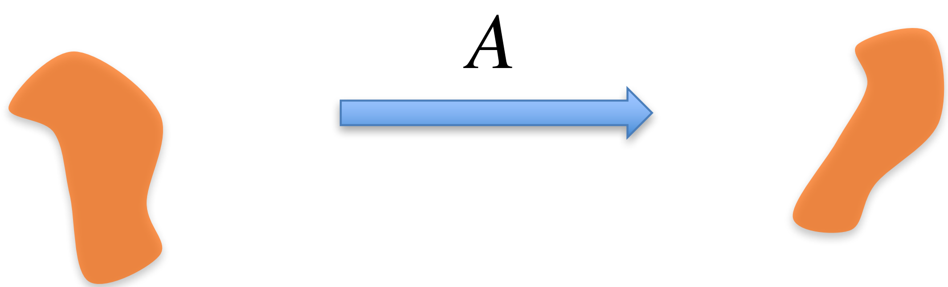
Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis

Matrix transformations

- In general A maps the set of vectors in $\mathbb{R}^2$ onto another set of vectors in $\mathbb{R}^2$.

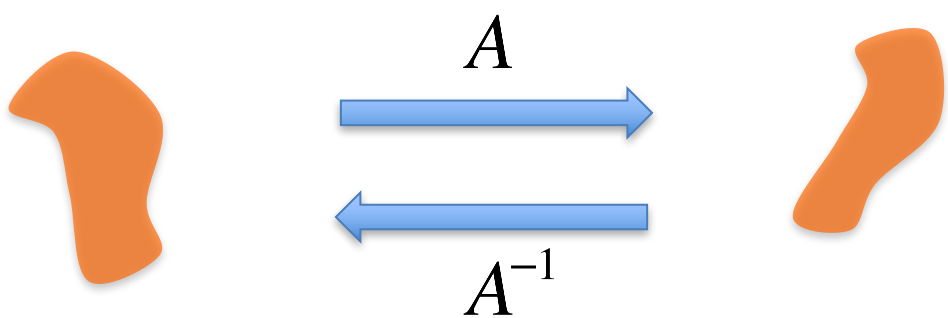
$$\vec{y} = A\vec{x}$$



Matrix transformations

- In general A maps the set of vectors in $\mathbb{R}^2$ onto another set of vectors in $\mathbb{R}^2$.

$$\vec{y} = A\vec{x}$$

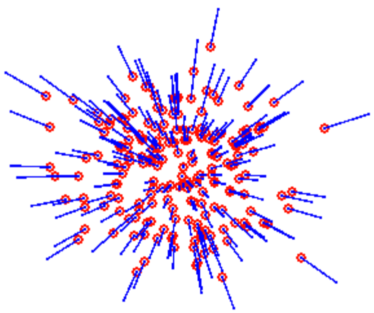


$$\vec{x} = A^{-1}\vec{y}$$

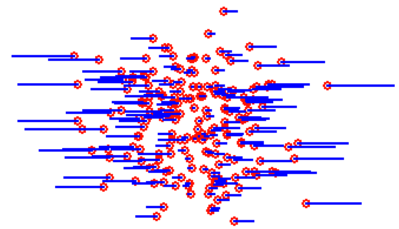
Matrix transformations $\vec{y} = A\vec{x}$

- Perturbations from the identity matrix

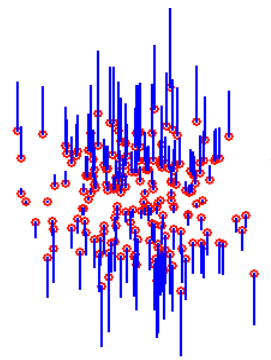
$$A = \begin{pmatrix} 1+\delta & 0 \\ 0 & 1+\delta \end{pmatrix}$$



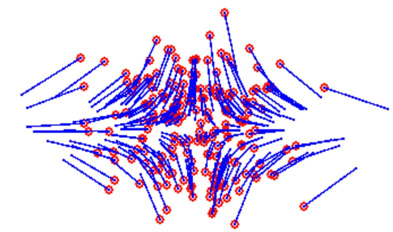
$$A = \begin{pmatrix} 1+\delta & 0 \\ 0 & 1 \end{pmatrix}$$



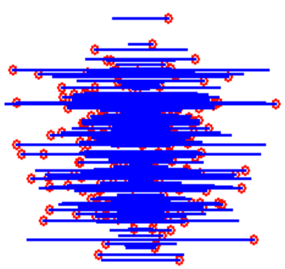
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1+\delta \end{pmatrix}$$



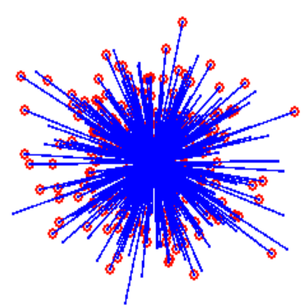
$$A = \begin{pmatrix} 1+\delta & 0 \\ 0 & 1-\delta \end{pmatrix}$$



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



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



These are all diagonal matrices

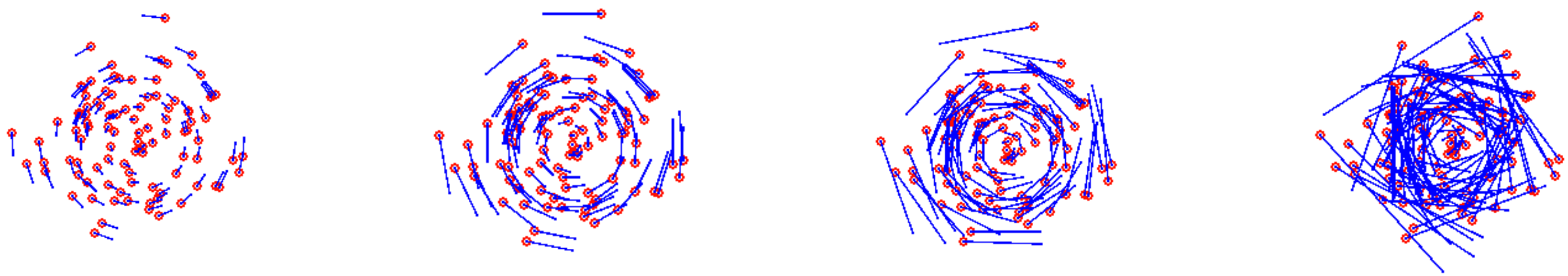
$$\Lambda = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

$$\Lambda^{-1} = \begin{pmatrix} a^{-1} & 0 \\ 0 & b^{-1} \end{pmatrix}$$

Rotation matrix

- Rotation in 2 dimensions $$\Phi(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

$\theta = 10^o$ $\theta = 25^o$ $\theta = 45^o$ $\theta = 90^o$



- Does a rotation matrix have an inverse? $\det(\Phi) = 1$

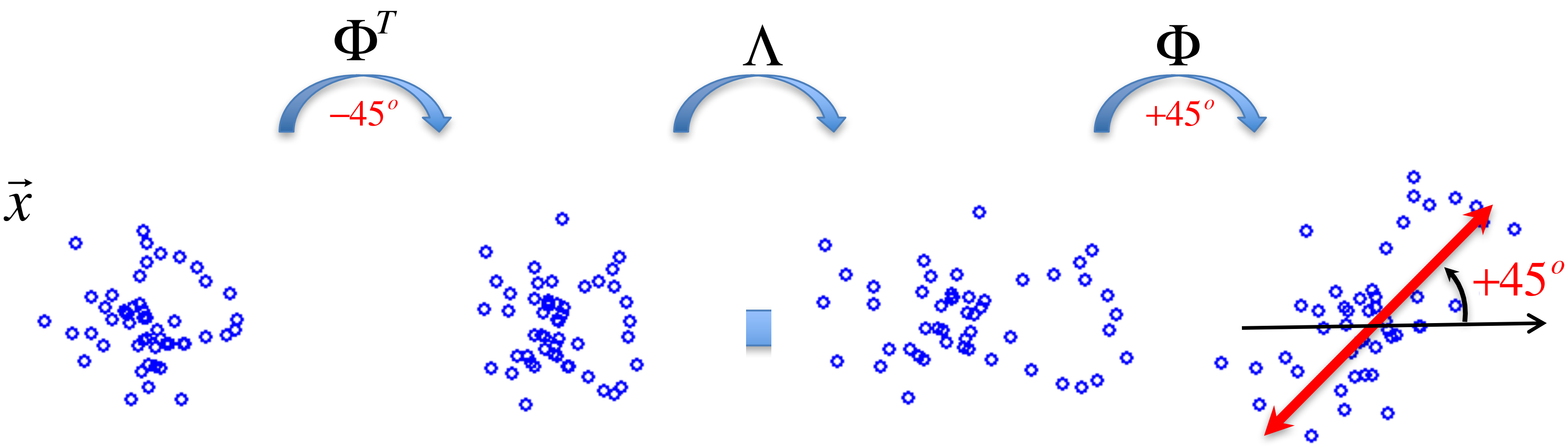
- The inverse of a rotation matrix is just its transpose

$$\Phi^{-1}(\theta) = \Phi(-\theta) = \Phi^T(\theta)$$

Rotated transformations

- Let's construct a matrix that produces a stretch along a 45° angle…

$$\Phi = \begin{pmatrix} \cos 45 & -\sin 45 \\ \sin 45 & \cos 45 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$



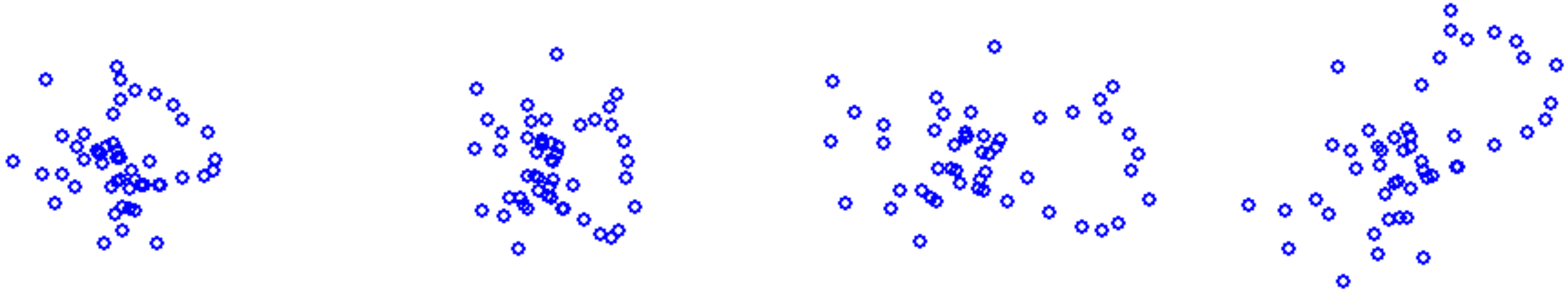
- We do each of these steps by multiplying our matrices together

$$\Phi \; \Lambda \; \Phi^T \; \vec{x}$$

Rotated transformations

- Let's construct a matrix that produces a stretch along a 45° angle…

$\vec{x}$ $\quad\quad$ $\Phi^T \vec{x}$ $\quad\quad$ $\Lambda\Phi^T \vec{x}$ $\quad\quad$ $\Phi\Lambda\Phi^T \vec{x}$



$$\Phi^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \qquad \Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$\Phi\Lambda\Phi^T = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Inverse of matrix products

- We can unto our transformation by taking the inverse

$$\left[\Phi\Lambda\Phi^T\right]^{-1}$$

- How do you take the inverse of a sequence of matrix multiplications A*B*C?

$$\left[ABC\right]^{-1} = C^{-1}B^{-1}A^{-1}$$

$$\begin{aligned}\left[ABC\right]^{-1}ABC &= C^{-1}B^{-1}A^{-1}ABC \\ &= C^{-1}B^{-1}BC \\ &= C^{-1}C \\ &= I\end{aligned}$$

- Thus…

$$\left[\Phi\Lambda\Phi^T\right]^{-1} = \left[\Phi^T\right]^{-1}\left[\Lambda\right]^{-1}\left[\Phi\right]^{-1}$$

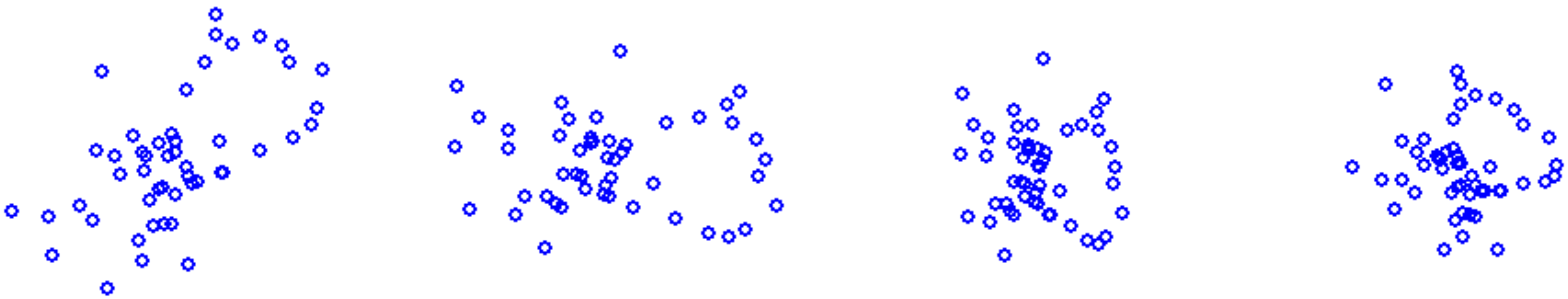
$$\left[\Phi\Lambda\Phi^T\right]^{-1} = \Phi\Lambda^{-1}\Phi^T$$

$$\Lambda^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix}$$

Rotated transformations

- Let's construct a matrix that undoes a stretch along a 45° angle…

$\vec{x}$ $\qquad \Phi^T \vec{x}$ $\qquad \Lambda^{-1}\Phi^T \vec{x}$ $\qquad \Phi\Lambda^{-1}\Phi^T \vec{x}$

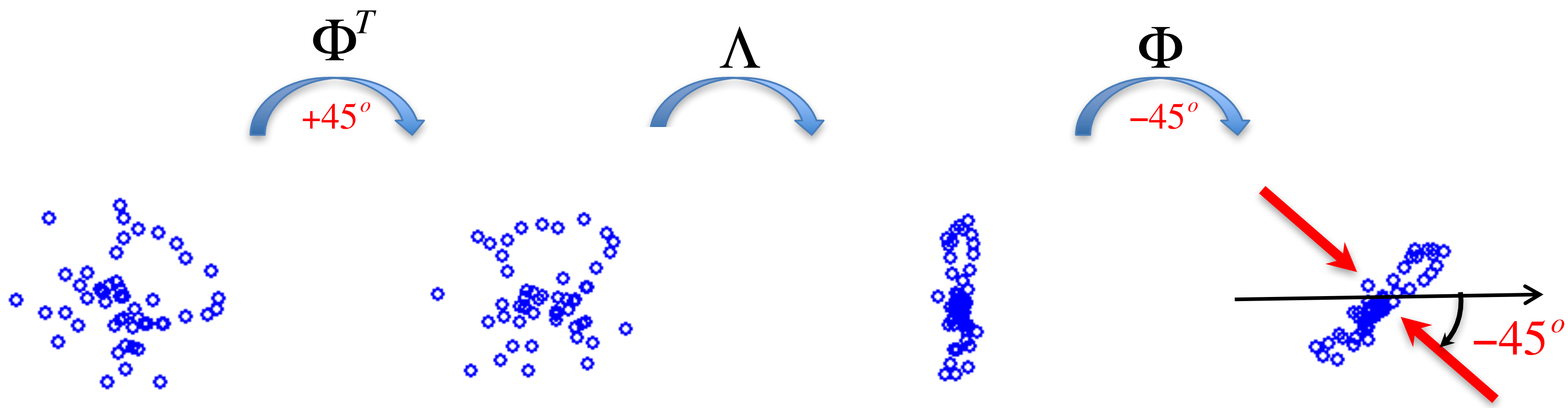


$$\Phi^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \qquad \Lambda^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Phi(+45^o) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$\Phi\Lambda^{-1}\Phi^T = \frac{1}{4}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

Rotated transformations

- Construct a matrix that does compression along a -45° angle…



$$\Phi^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} 0.2 & 0 \\ 0 & 1 \end{pmatrix}$$

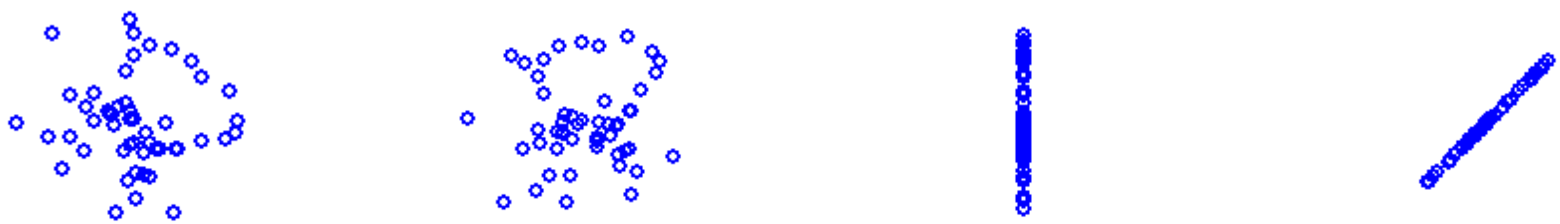
$$\Phi(-45^o) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$\Phi\Lambda\Phi^T = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.4 \end{pmatrix}$$

Transformations that can't be undone

- Some transformation matrices have no inverse…

$\vec{x}$ $\qquad \Phi^T \vec{x}$ $\qquad \Lambda\Phi^T \vec{x}$ $\qquad \Phi\Lambda\Phi^T \vec{x}$



$$\Phi^T = \Phi(45^o) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad \Lambda = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Phi(-45^o) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$\Phi\Lambda\Phi^T = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \qquad \det\left(\Phi\Lambda\Phi^T\right) = 0 \qquad \det(\Lambda) = 0$$

Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis

Basics of basis sets

- We can think of vectors as abstract 'directions' in space. But in order to specify the elements of a vector, we need to choose a coordinate system.

- To do this, we write our vector as a linear combination of a set of special vectors called the 'basis set.'

$$\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x\hat{e}_1 + y\hat{e}_2 + z\hat{e}_3$$

- The numbers x, y, z are called the coordinates of the vector.

- The vectors $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ are called the 'basis vectors', in this case, in three dimensions.

Basics of basis sets

- In order to describe an arbitrary vector in the space of real numbers in n dimensions ($\mathbb{R}^n$), our basis vectors need to have n numbers.

- In order to describe an arbitrary vector in $\mathbb{R}^n$, we need to have n basis vectors.

- The basis set we wrote down earlier $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ is called the '<u>standard basis</u>'. Each vector has one element that's a one and the rest are zeros.

$$\hat{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \hat{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Orthonormal basis

- In addition, the standard basis has the interesting property that

  each vector is a unit vector

$$\hat{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \hat{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\hat{e}_i \cdot \hat{e}_i = 1$$

- Each vector is orthogonal to all the other vectors

$$\hat{e}_1 \cdot \hat{e}_2 = 0 \qquad \hat{e}_1 \cdot \hat{e}_3 = 0 \qquad \hat{e}_2 \cdot \hat{e}_3 = 0 \qquad \hat{e}_i \cdot \hat{e}_j = 0\,, \quad i \neq j$$

- These properties can be written compactly as

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} \qquad \delta_{ij} = \begin{cases} 1 & if \;\; i = j \\ 0 & if \;\; i \neq j \end{cases}$$

- Any basis set with these properties is called 'orthonormal'.

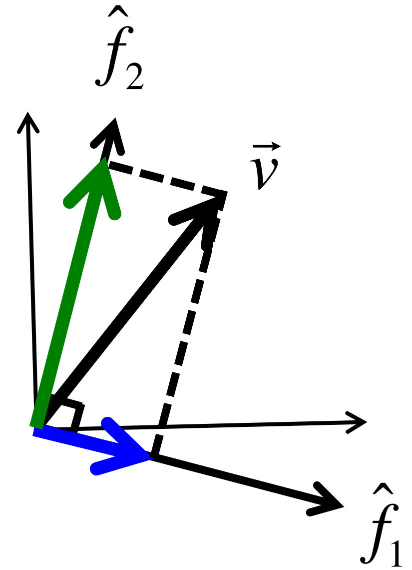
Basics of basis sets

- The standard basis is not the only orthonormal basis

  Consider a different set of orthogonal unit vectors: $\{\vec{f}_1, \vec{f}_2\}$

$$\vec{v} = \left(\vec{v}\cdot\hat{f}_1\right)\hat{f}_1 + \left(\vec{v}\cdot\hat{f}_2\right)\hat{f}_2$$

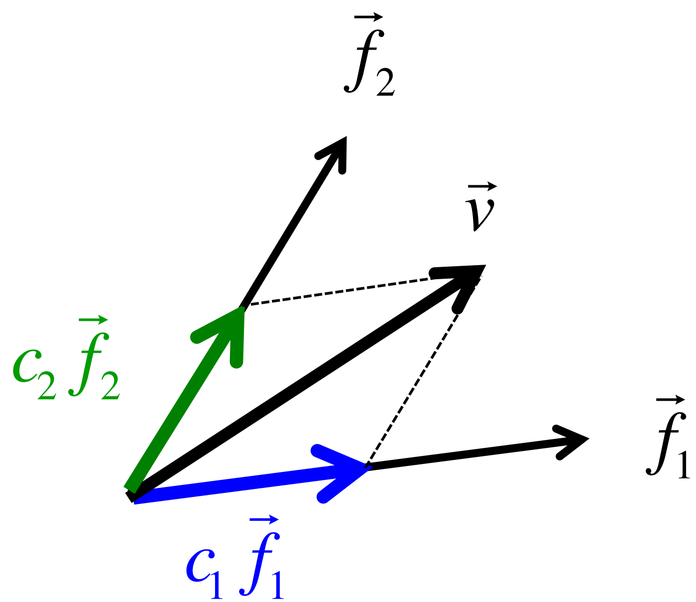
$$\vec{v}_f = \begin{pmatrix} \vec{v}\cdot\hat{f}_1 \\ \vec{v}\cdot\hat{f}_2 \end{pmatrix}$$



- The vector coordinates are given by the dot products of the vector $\vec{v}$ with each of the basis vectors.

Non-orthonormal basis sets

- Vectors can also be written as a linear combination of (almost) any vectors, not just orthonormal basis vectors



$$\vec{v} = c_1\vec{f}_1 + c_2\vec{f}_2$$

Basics of basis sets

- Let's decompose an arbitrary vector v in a basis set $\{\vec{f}_1, \vec{f}_2\}$

$$\vec{v} = c_1 \vec{f}_1 + c_2 \vec{f}_2$$

- The coefficients c_1 and c_2 are called 'coordinates of the vector v in the basis $\{\vec{f}_1, \vec{f}_2\}$.

- The vector $\vec{v}_f = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ is called the 'coordinate vector' of $\vec{v}$

  in the basis $\{\vec{f}_1, \vec{f}_2\}$.

Basics of basis sets

- Let's look at an example. Consider the basis

$$\{\vec{f}_1, \vec{f}_2\} = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$$

and the vector $\vec{v} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ in the standard basis.

- Find the vector coordinates of $\vec{v}$ in the new basis.

- Write $\vec{v}$ as a linear combination of the new basis vectors:

$$c_1 \vec{f}_1 + c_2 \vec{f}_2 = \vec{v}$$

$$c_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

system of equations

$$c_1 - 2c_2 = 3$$
$$3c_1 + c_2 = 5$$

Basics of basis sets

- We can write this system of equations in matrix notation:

$$c_1 - 2c_2 = 3$$
$$3c_1 + c_2 = 5$$

$$F\,\vec{v}_f = \vec{v}$$

where $F = \begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}$ $\quad \vec{v}_f = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ $\quad \vec{v} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$

- Now solve for $\vec{v}_f$ by multiplying both sides of the equation by the inverse of matrix F .

$$F^{-1}\,F\,\vec{v}_f = F^{-1}\vec{v}$$

$$\vec{v}_f = F^{-1}\,\vec{v}$$

Basics of basis sets

- We can find the inverse of F :

$$F^{-1} = \frac{1}{7}\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}$$

$$\vec{v}_f = F^{-1}\vec{v} \quad = \frac{1}{7}\begin{pmatrix} 1 & 2 \\ -3 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$= \frac{1}{7}\begin{pmatrix} 3+10 \\ -9+5 \end{pmatrix} \quad = \frac{1}{7}\begin{pmatrix} 13 \\ -4 \end{pmatrix}$$

- Thus, we find the coordinate vector of v in basis $\{\vec{f}_1, \vec{f}_2\}$

$$\vec{v}_f = \begin{pmatrix} 13/7 \\ -4/7 \end{pmatrix}$$

Basics of basis sets

- In summary: to find the coordinate vector for v in the basis $\{\vec{f}_1, \vec{f}_2\}$, we construct a matrix F whose columns are just the elements of the basis vectors.

$$F = \left(\vec{f}_1 \middle| \vec{f}_2 \right) \qquad\qquad F = \left(\vec{f}_1 \middle| \vec{f}_2 \middle| \vec{f}_3 \ldots \middle| \vec{f}_n \right)$$

such that $\vec{v} = F\,\vec{v}_f$

- We can solve for $\vec{v}_f$ by multiplying both sides of the equation by the inverse of matrix F

$$\vec{v}_f = F^{-1}\,\vec{v}$$

'change of basis'

- But this only works if F has an inverse!

Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis

Subspaces

- We need n vectors in $\mathbb{R}^n$ to form a basis in $\mathbb{R}^n$. But not any set of n vectors will do the trick!

- Consider the following set of vectors

$$\vec{f}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \vec{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \vec{f}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

- Note that any linear combination of $\{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ will always lie in the (x , y) plane

$$\vec{v} = c_1 \vec{f}_1 + c_2 \vec{f}_2 + c_3 \vec{f}_3 \qquad = \begin{pmatrix} c_1 + c_3 \\ c_2 + c_3 \\ 0 \end{pmatrix}$$

- Thus, the set of vectors $\{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ doesn't span all of $\mathbb{R}^3$

  It only spans the x-y plane - a subspace of $\mathbb{R}^3$

Linear independence

$$\vec{f}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \vec{f}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \vec{f}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

- Note that we can write any of these vectors as a linear combination of the other two.

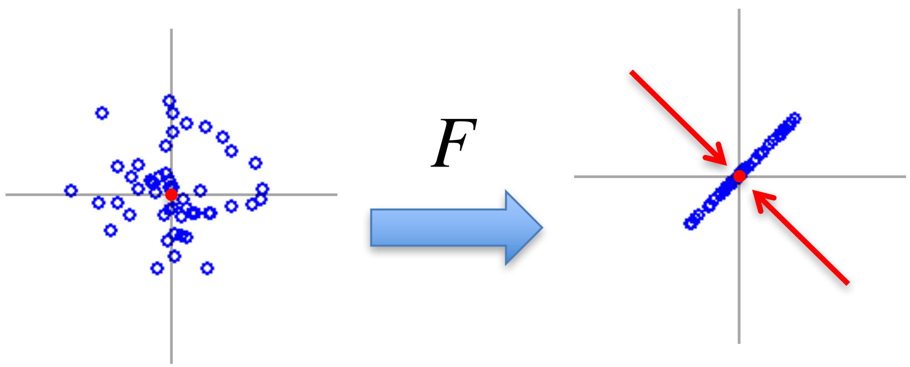
$$\vec{f}_3 = \vec{f}_1 + \vec{f}_2 \qquad \vec{f}_2 = \vec{f}_3 - \vec{f}_1 \qquad \vec{f}_1 = \vec{f}_3 - \vec{f}_2$$

- Thus, this set of vectors is called 'linearly dependent'.
- Any set of n linearly dependent vectors cannot form a basis in $\mathbb{R}^n$
- How do you know if a set of vectors is linearly dependent?

$$F = \left(\vec{f}_1 \middle| \vec{f}_2 \middle| \vec{f}_3 \ldots \middle| \vec{f}_n \right) \qquad \det(F) = 0$$

Linear dependence

- If $\det(F) = 0$ then F maps $\vec{v}_f$ into a subspace of $\mathbb{R}^n$



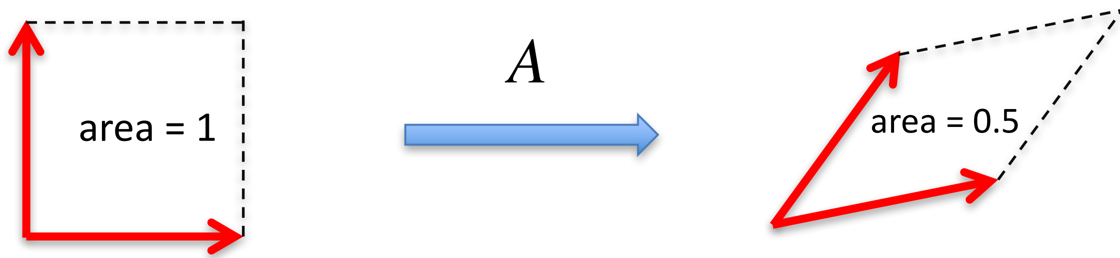
$$F = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

- If F maps onto a subspace, then the mapping is not reversible!

$$\det(F) = 0$$

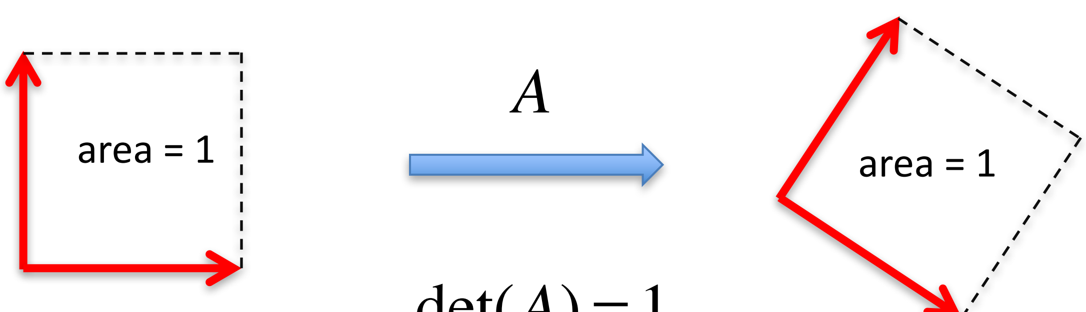
Geometric interpretation of determinant

- The determinant is the ‘volume’ of a unit cube after transformation (area of unit square in two dimensions).



$$\det(A) = 0.5$$

- A pure rotation matrix has a determinant of one.



$$\det(A) = 1$$

Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis

Change of basis

$$\left\{\vec{f}_1, \vec{f}_2, \dots \vec{f}_n\right\} \qquad F = \left(\vec{f}_1 \,\middle|\, \vec{f}_2 \,\middle|\, \dots \,\middle|\, \vec{f}_n\right)$$

- If $\det(F) \neq 0$ then the vectors $\left\{\vec{f}_1, \vec{f}_2, \dots \vec{f}_n\right\}$
  - are linearly independent
  - form a complete basis set in $\mathbb{R}^n$

- Then the matrix F implements a 'change of basis'

From standard basis to $\vec{f}$

$$\vec{v}_f = F^{-1}\vec{v}$$

Or from $\vec{f}$ to standard basis

$$\vec{v} = F\,\vec{v}_f$$

Change of basis

- The change of basis is easy if $\{\vec{f}_1, \vec{f}_2\}$ is an orthonormal basis…

$$F = \begin{pmatrix} | & | \\ \hat{f}_1 & \hat{f}_2 \\ | & | \end{pmatrix} \qquad F^T = \begin{pmatrix} - & \hat{f}_1 & - \\ - & \hat{f}_2 & - \end{pmatrix}$$

$$F^T F = \begin{pmatrix} - & \hat{f}_1 & - \\ - & \hat{f}_2 & - \end{pmatrix} \begin{pmatrix} | & | \\ \hat{f}_1 & \hat{f}_2 \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Thus… $F^T = F^{-1}$

F is just a rotation matrix!

Change of basis

- With an orthonormal basis set, the coordinates are just given by the dot product with the basis vectors !

$$F = \begin{pmatrix} | & | \\ \hat{f}_1 & \hat{f}_2 \\ | & | \end{pmatrix} \qquad F^{-1} = F^T = \begin{pmatrix} - & \hat{f}_1 & - \\ - & \hat{f}_2 & - \end{pmatrix}$$

$$\vec{v}_f = F^{-1}\vec{v}$$

$$\vec{v}_f = F^T\vec{v} = \begin{pmatrix} - & \hat{f}_1 & - \\ - & \hat{f}_2 & - \end{pmatrix}\vec{v} = \begin{pmatrix} \vec{v}\cdot\hat{f}_1 \\ \vec{v}\cdot\hat{f}_2 \end{pmatrix}$$

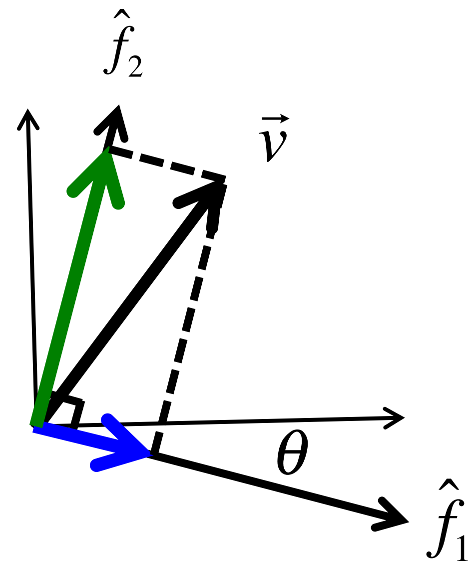
Change of basis

- In two dimensions, there is a family of orthonormal basis sets

$$\hat{f}_1 = \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix} \qquad \hat{f}_2 = \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix} \qquad F = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

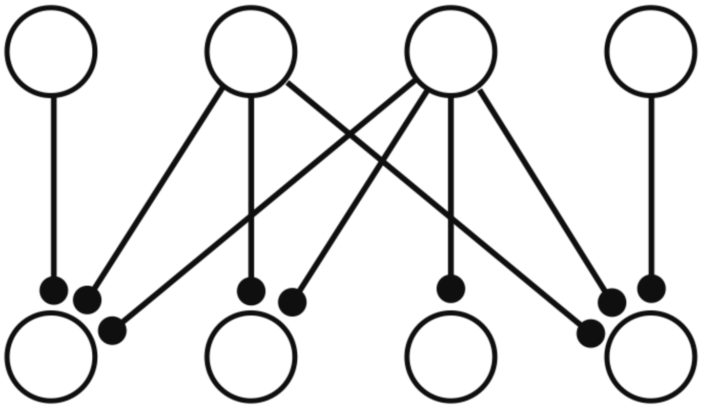
$$\vec{v} = \left(\vec{v}\cdot\hat{f}_1\right)\hat{f}_1 + \left(\vec{v}\cdot\hat{f}_2\right)\hat{f}_2$$

$$\vec{v}_f = F^T\vec{v} \qquad \vec{v}_f = \begin{pmatrix} \vec{v}\cdot\hat{f}_1 \\ \vec{v}\cdot\hat{f}_2 \end{pmatrix}$$



- The vector coordinates are given by the dot products of the vector $\vec{v}$ with each of the rotated basis vectors.

Seeing in high dimensions



https://research.googleblog.com/2016/12/open-sourcing-embedding-projector-tool.html

Learning Objectives for Lecture 16

- More on two-layer feed-forward networks
- Matrix transformations (rotated transformations)
- Basis sets
- Linear independence
- Change of basis