Reinforcement Learning

Lecture 1

Course Organization

Intro to Reinforcement Learning

Mathematical Basics

Robert Peharz

Institute of Theoretical Computer Science
Graz University of Technology
Winter Term 2022/23

Course Organization

Robert Peharz

- Assistant Professor at TU Graz
- Institute of Theoretical Computer Science (IGI)
- Machine Learning, Artificial Intelligence
- short bio
  - studied here (Telematics, PhD)
  - Meduni Graz (2015-2017)
  - Cambridge, UK (2017-2019)
  - TU Eindhoven, NL (2019-2021)
  - TU Graz, since October 2021

robert.peharz@tugraz.at

Lectures

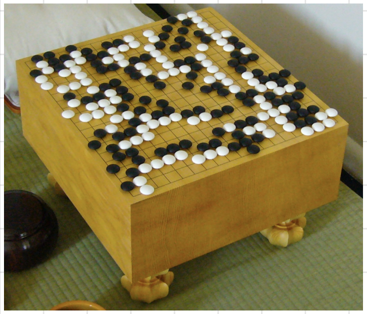
2 x 45 min. + 10 min. break

708.061 Reinforcement Learning (2SSt VO, WS 2023/24)

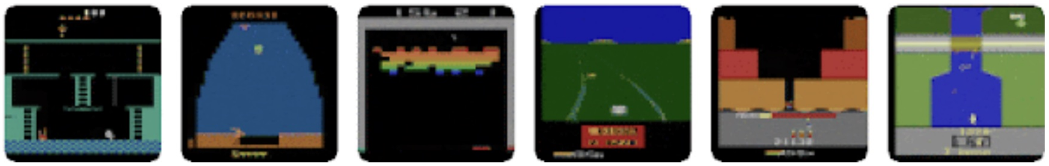
Gruppe						
Tag	Datum	von	bis	Ort	Ereignis	Termintyp
Standardgruppe						
Mi	04.10.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
~~Mi~~	~~11.10.2023~~	~~09:45~~	~~11:30~~	~~HS i11 "SIEMENS Hörsaal" (ICK1002H)~~	~~Abhaltung~~	~~fix~~
~~Mi~~	~~18.10.2023~~	~~09:45~~	~~11:30~~	~~HS i11 "SIEMENS Hörsaal" (ICK1002H)~~	~~Abhaltung~~	~~fix~~
Mi	25.10.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	08.11.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	15.11.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	22.11.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	29.11.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	06.12.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	13.12.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	20.12.2023	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	10.01.2024	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	17.01.2024	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix
Mi	24.01.2024	09:45	11:30	HS i11 "SIEMENS Hörsaal" (ICK1002H)	Abhaltung	fix

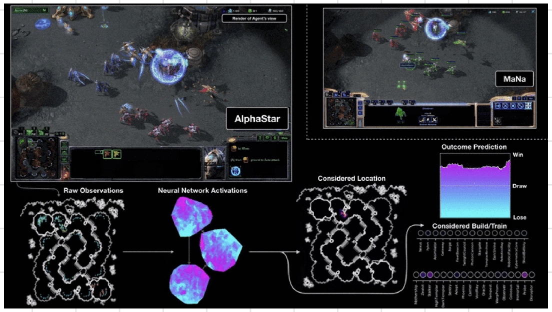
Reinforcement Learning

- one of the big branches in machine learning
- hot topic
  - mastering the Game of Go

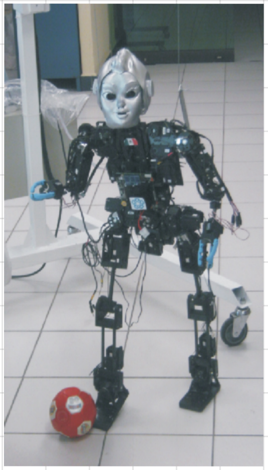


  - Learning to play Computer Games





  - teaching robots to walk



  - ...

Images: wikipedia.org, deepmind.com, Katic et al. 2012

Lecture (VO)

- theory and foundations of RL
- algorithms
- examples and applications
- discussing research papers

Learning Goals

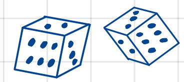
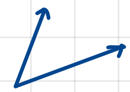
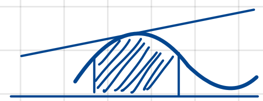
- getting started in RL
- prepare for RL projects (Master, PhD, industry)
- augment your ML tool kit

Practicals (KU)
Fr. 12:30-14:00, i11

Özdenizci & Wedenig

- 2-3 Homework Assignments
- implement RL algorithms
- some theory exercises
- reading assignments + presentation

What do you need for this course?

- solid understanding of basic
  - probability theory 
  - linear algebra 
  - calculus 
- ideally, an introductory machine learning course (e.g. ML1)
- python skills, ideally pyTorch (for the KU)

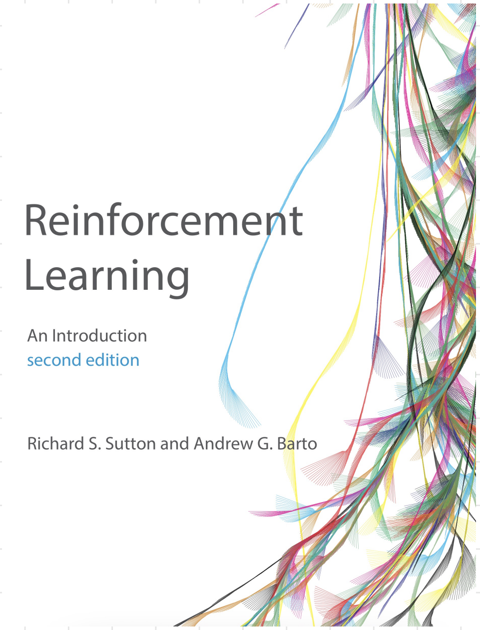
Exam

- written exam of 90 minutes
- open questions and/or multiple choice, checking understanding and analytic skills
- small computational examples checking ability to apply concepts and algorithms (pocket calculator needed)
- 100 points

> 90	Sehr Gut	(1)
75.1 - 90	Gut	(2)
65.1 - 75	Befriedigend	(3)
50.1 - 65	Genügend	(4)
0 - 50	Nicht Genügend	(5)

next exam: Thur. 12th Oct, 10:00 , 11 ; then end of semester

Relevant Materials



- accesible introduction
- written by two pioneers of modern RL
- freely available

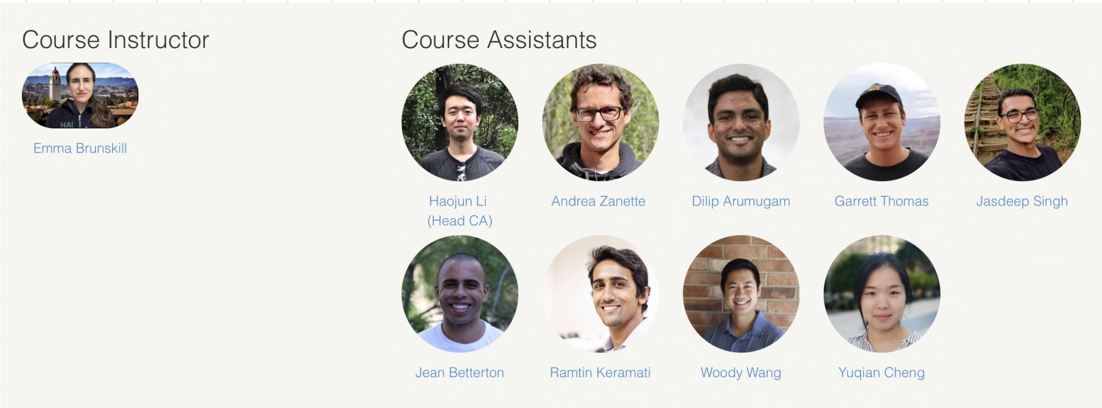
https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

- excellent video lectures
- by a leading RL researcher
- 10 lectures of 1.5 hours
- slides & videos freely available

https://deepmind.com/learning-resources/-introduction-reinforcement-learning-david-silver

More Materials

Stanford RL course



https://web.stanford.edu/class/cs234/

Algorithms for Reinforcement Learning, Szepesvari

- short and concise book
- more of a "mathematical style"

https://sites.ualberta.ca/~szepesva/rlbook.html

Contact

- Teach Center, Forum
- robert.peharz@tugraz.at

Color Codes in the Notes

- narrator's voice
- math
- definition, introducing notions
- example
- recap, summary
- important note, question to think about
- algorithm (as font color)
- side note

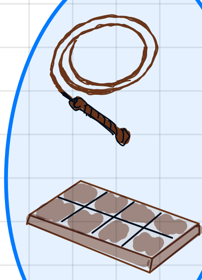
Not 100% consistent

And sometimes colors are used as (hopefully) visual aids

Introduction: What is Reinforcement Learning?

Reinforcement Learning is about...

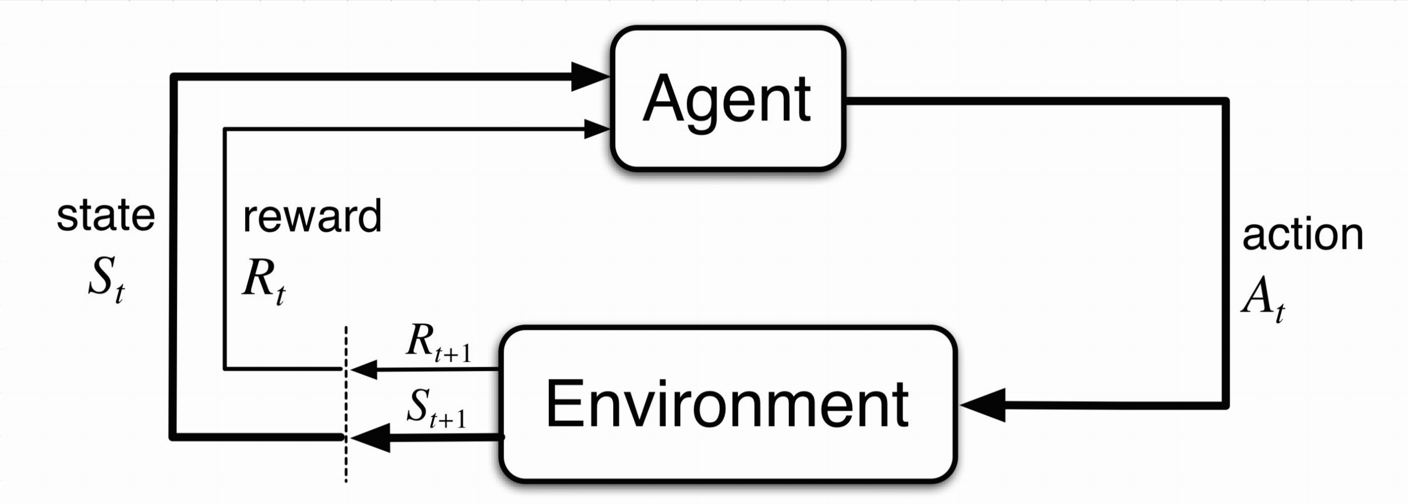
- an **agent** interacting with an **environment**
- agent makes decisions (picks **actions**) over time
  → sequential decision making
- agent's actions influence environment
  - environment changes its **state**
  - environment returns a **reward**
    - "low ≈ pain"
    - "high ≈ pleasure"



- goal: maximize reward over agent's lifetime
- learn behavior (**policy**) from weak feedback
- behavior leading to high reward gets reinforced
  (term from psychology)

Agent-Environment Interface

Image: Sutton & Barto



- discrete time (usually, and so in this course)
- at time t, environment is in state S_t
- agent plays action A_t
- based on S_t and A_t, environment
  - switches to state S_{t+1}
  - emits reward R_{t+1}

  observed by agent
- agent selects next action A_{t+1} (and so on...)

Agent vs. Environment

- what belongs to agent, what to environment?
- boundary is often closer to the agent than expected
- "anything which cannot be changed completely arbitrarily should be seen as part of the environment" (Sutton & Barto)
- boundary of absolute control (not of knowledge)
- e.g.
  - position of a robot,
  - angles of your limbs,
  - etc.

  should be seen as part of the environment

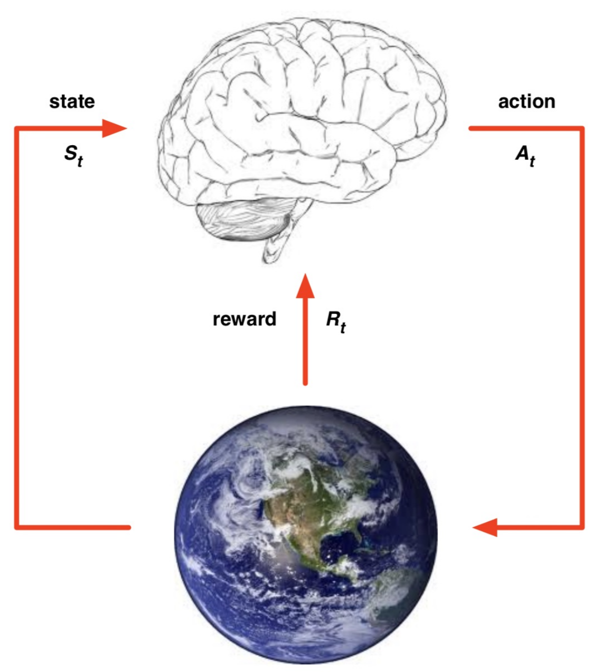


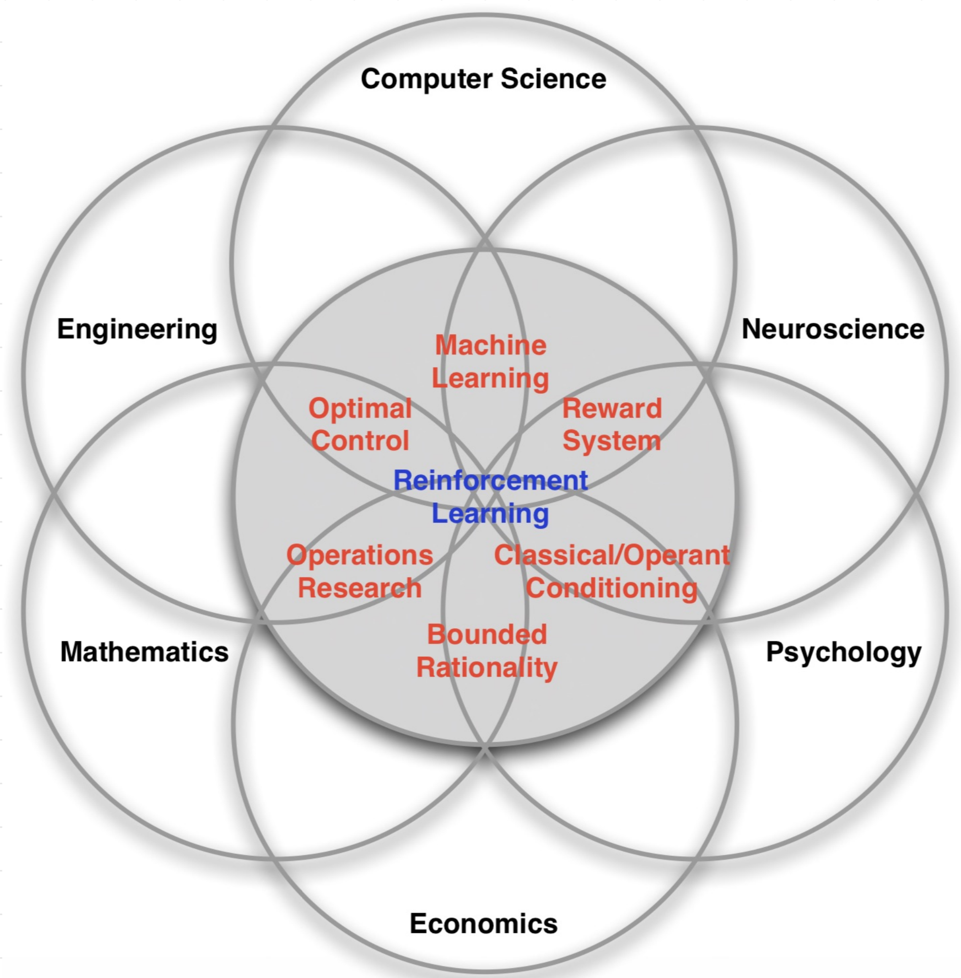
Image: D. Silver

Examples of Reinforcement Learning

- cleaning robot
- battery control system in same robot (nested RL)
- controler in a petroleum refinery
- new-born gazelle : barely able to stand, but runs 20 Km/h after 30 min.
- advertisement placement (websites)
- portfolio optimization

What would be good reward signals ?

The Many Facets of Reinforcement Learning



Reinforcement Learning: the science of decision making

Image: D. Silver

RL as Third Pillar of Machine Learning (?)

- RL is not supervised learning
  (only rewards as "weak feedback", but no expert labels)
- RL is not unsupervised learning
  (finding structure in data might help, but is not a goal per se)
- RL is interactive (agent needs to explore the environment) and sequential (i.e. inherently time-dependent).
  These are not typical features of supervised/unsupervised learning.

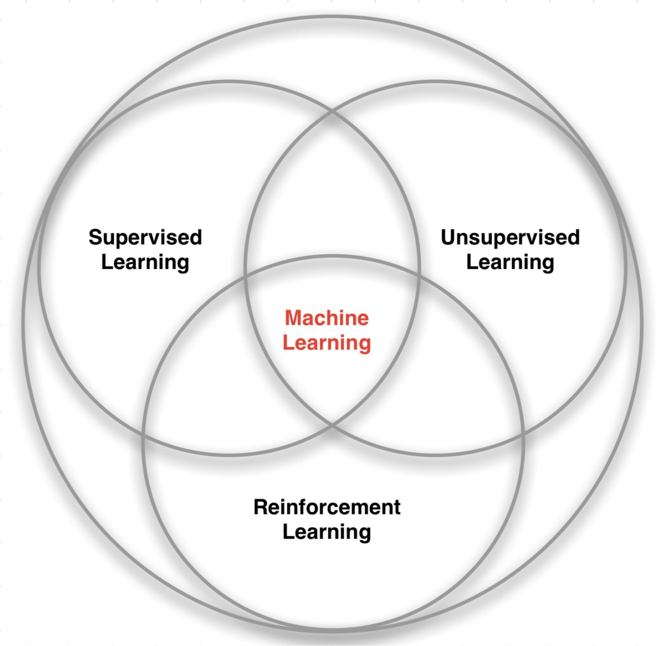


Image: D. Silver

There are, however, many other ML paradigms, which do have these characteristics as well: !

active learning, semi-supervised learning, Bayesian optimization, coactive learning, ...

Atari Breakout

https://www.youtube.com/watch?v=TmPfTpjtdgg

Go

https://www.youtube.com/watch?v=7L2sUGcOgh0

Star Craft

https://www.youtube.com/watch?v=UuhECwm31dM

Cart Pole

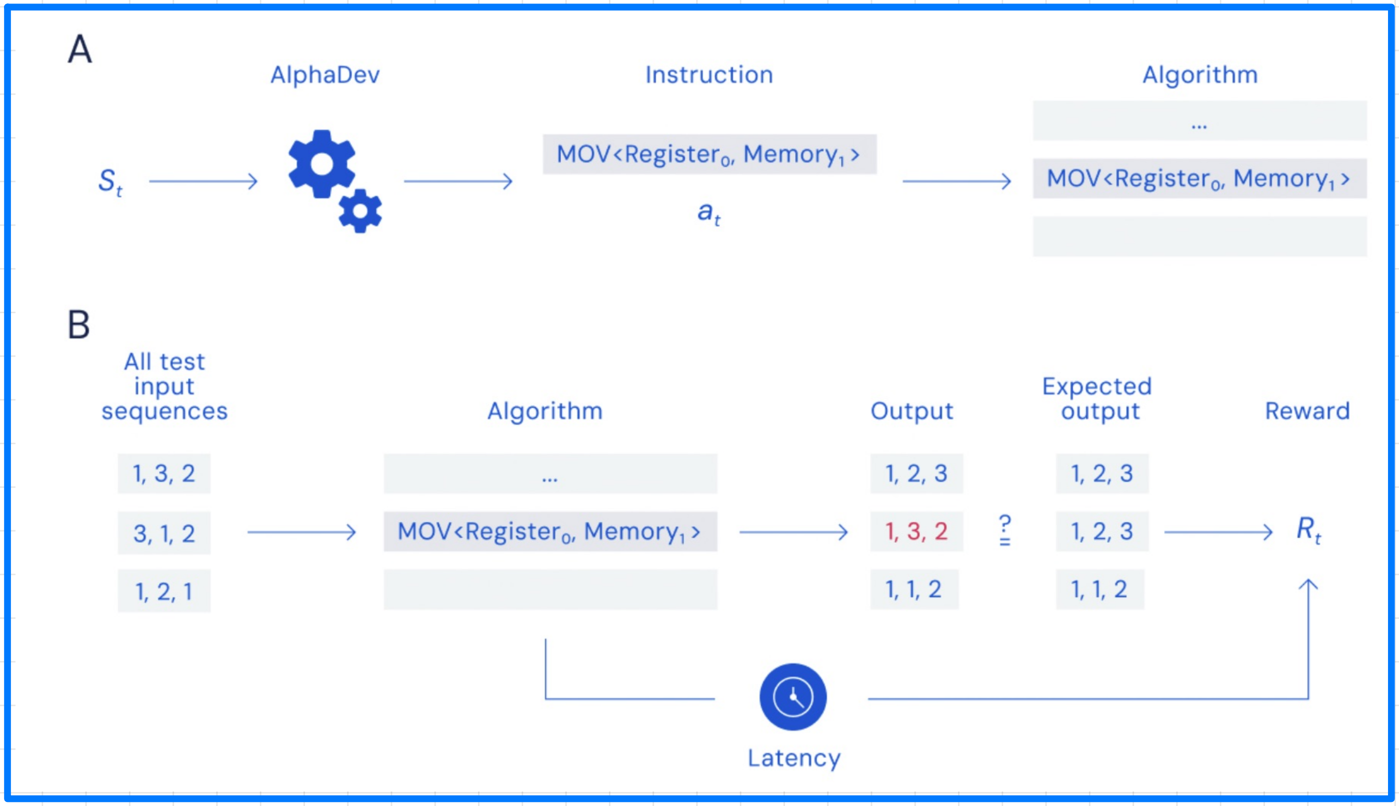
http://www.youtube.com/watch?v=XiigTGKZfks

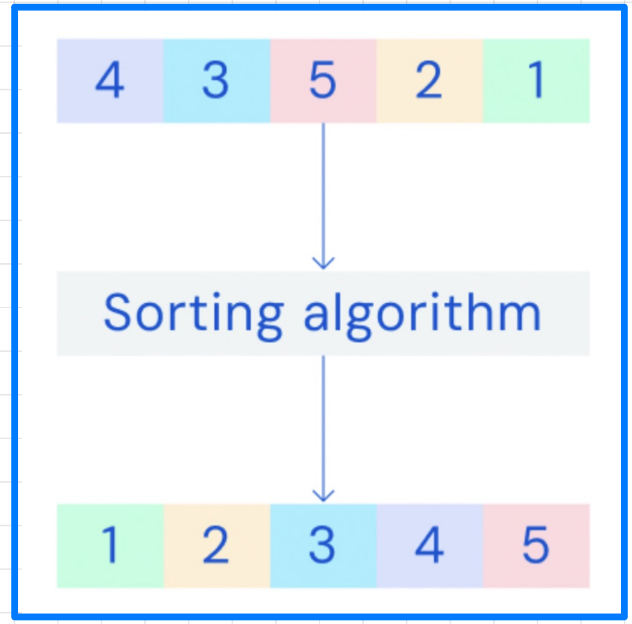
Robot Learning to Walk

https://www.youtube.com/watch?v=goxCjGPQH7U

Discovering New Algortihms

Mankowitz, D.J., Michi, A., Zhernov, A. *et al.* Faster sorting algorithms discovered using deep reinforcement learning. *Nature* **618**, 257–263 (2023).





Original

```
Memory[0] = A
Memory[1] = B
Memory[2] = C

mov Memory[0] P  // P = A
mov Memory[1] Q  // Q = B
mov Memory[2] R  // R = C

mov R S
cmp P R
cmovg P R  // R = max(A, C)
cmovl P S  // S = min(A, C)
mov S P    // P = min(A, C)
cmp S Q
cmovg Q P // P = min(A, B, C)
cmovg S Q // Q = max(min(A, C), B)

mov P Memory[0]  // = min(A, B, C)
mov Q Memory[1]  // = max(min(A, C), B)
mov R Memory[2]  // = max(A, C)
```

AlphaDev

```
Memory[0] = A
Memory[1] = B
Memory[2] = C

mov Memory[0] P  // P = A
mov Memory[1] Q  // Q = B
mov Memory[2] R  // R = C

mov R S
cmp P R
cmovg P R  // R = max(A, C)
cmovl P S  // S = min(A, C)

cmp S Q
cmovg Q P  // P = min(A, B)
cmovg S Q  // Q = max(min(A, C), B)

mov P Memory[0]  // = min(A, B)
mov Q Memory[1]  // = max(min(A, C), B)
mov R Memory[2]  // = max(A, C)
```

Left: The original implementation with min(A,B,C).
Right: AlphaDev Swap Move – AlphaDev discovers that you only need min(A,B).

- 70% faster for short sequences
- 1.7% faster for long sequences
- new "moves"

Recap: Probability

Notation for Random Variables

- random variables (RVs) are denoted by upper-case letters

  X, Y, Z, S_t, A_t, R_t

- states of RVs are denoted by lower-case letters

  x, y, z, s, a, r

- state spaces of RVs (set of all states) as calligraphic letters

  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$

- underline for random vectors (multivariate RVs)

  $\underline{X}, \underline{Y}, \underline{Z}$

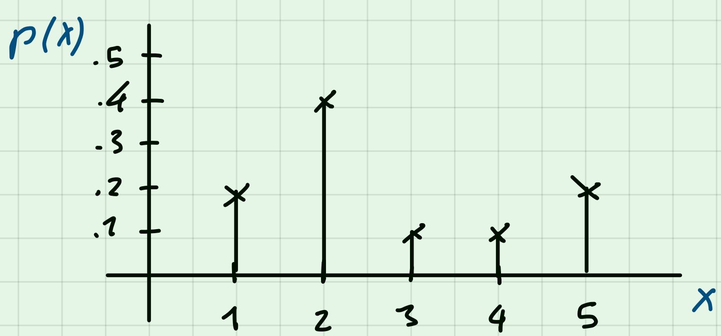
Discrete Random Variable, Probability Mass Function

Let X be a random variable with state space $\mathcal{X}$ (i.e. the set of values X can assume). If $\mathcal{X}$ is countable then X is called a **discrete** random variable.

countable set: finite or of same cardinality as $\mathbb{N}$. E.g. $\{1,2,3\}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$

Let p_X be the function $p_X : \mathcal{X} \to [0,1]$ with $p_X(x) = \mathbb{P}(X = x)$, i.e. the function which assigns each state $x \in \mathcal{X}$ the probability that $X = x$. This function p_X is called the **probability mass function (pmf)** of X.

Example pmf of a random variable X with $\mathcal{X} = \{1,2,3,4,5\}$.



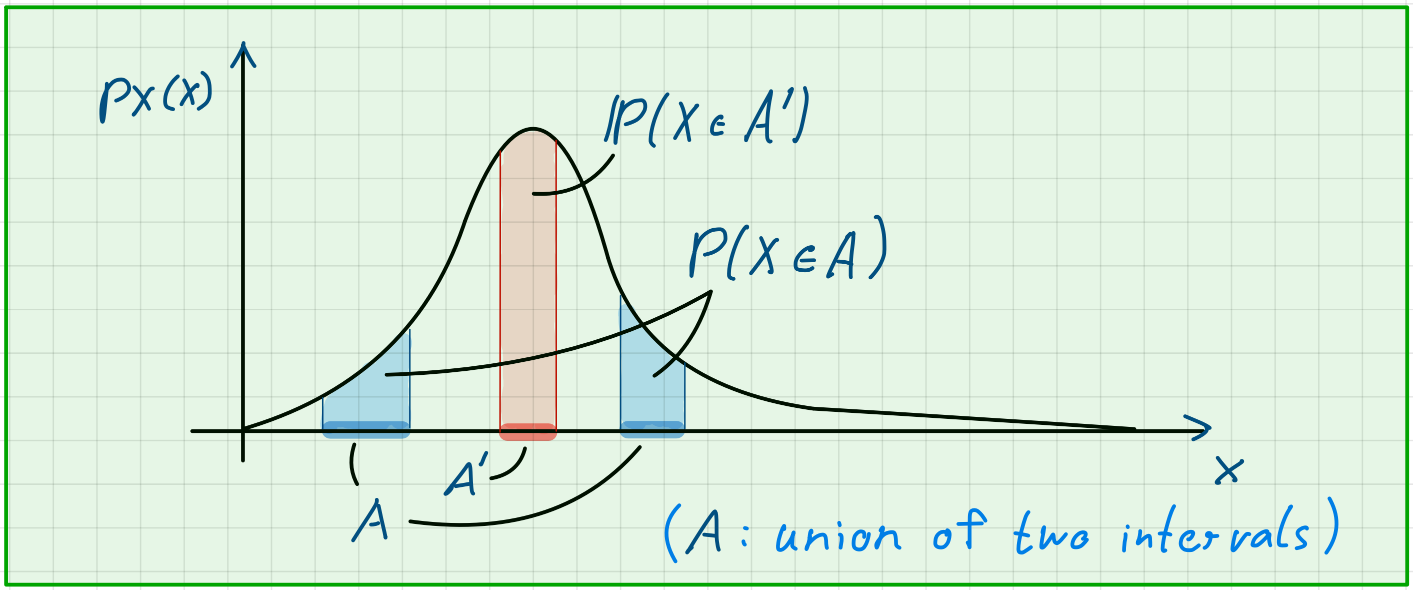
Continuous Random Variable, Probability Density

Let X be a random variable with state space $\mathcal{X} = \mathbb{R}$. If there exists a function $p_X : \mathcal{X} \to \mathbb{R}_+$ ($\mathbb{R}_+$ are the non-negative real numbers) with

$$\int_A p_X(x)\, dx = \mathbb{P}(X \in A)$$

then p_X is called a **probability density function (pdf)** (or just **density**) of X. Often we omit the subscript and write p.

A random variable X which has a density is called **continuous**.



Multivariate Random Variable, Joint Distributions

Both pmfs and pdfs generalize to multivariate random variables (random vectors) $\underline{X} = (X_1, X_2, \ldots, X_D)$. Here, X_d, $d = 1..D$, are random variables combined to a random vector $\underline{X}$.

Joint pmf

$$p_{\underline{X}}(x_1, x_2, \ldots, x_D) := \mathbb{P}(\underline{X} = (x_1, x_2, \ldots, x_D))$$

Joint pdf

$$\int_A p_{\underline{x}}(x_1, x_2, \ldots, x_D)\, d\underline{x} = \mathbb{P}(\underline{X} \in A) \qquad A \subseteq \mathbb{R}^D$$

- sloppy notation: $p(x)$ to mean either p_X or $p_X(x)$, i.e. getting rid of subscript
- use the same symbol p for pmfs and pdfs. We do this since both have the same purpose: compact representation of a distribution.

Marginal Distribution

Assume random variables $\underline{X} = \{X_1, X_2, \dots, X_D\}$ with joint distribution
$p_{\underline{X}}(x_1, x_2, \dots, x_D)$

Split $\{X_1, X_2, \dots, X_D\}$ into two arbitrary disjoint sets

$\underline{Y} = \{Y_1, Y_2, \dots, Y_I\}$ and $\underline{Z} = \{Z_1, Z_2, \dots, Z_J\}$, $\underline{Y} \cup \underline{Z} = \underline{X}$, $\underline{Y} \cap \underline{Z} = \{\}$.

The **marginal** distribution over $\underline{Y}$ is given as

$$p_{\underline{Y}}(y_1, y_2, \dots, y_I) = \sum_{z_1} \sum_{z_2} \dots \sum_{z_J} p_{\underline{X}}(y_1, y_2, \dots, y_I, z_1, z_2, \dots, z_J)$$

- for continuous random varibles replace $\sum_{z_j}$ with $\int \dots dz_j$.
- $p_{\underline{X}}$: describes all random variables
- $p_{\underline{Y}}$: describes $\underline{Y} \subseteq \underline{X}$, ignoring (accounting for) $\underline{Z} = \underline{X} / \underline{Y}$

Example: Marginal Distribution

Assume 3 binary random variables X, Y, Z and let the joint pmf be

x	y	z	$p(x,y,z)$
0	0	0	0.1
0	0	1	0.15
0	1	0	0.05
0	1	1	0.2
1	0	0	0.1
1	0	1	0.1
1	1	0	0.15
1	1	1	0.15

marginalizing Y

The pmf $p(x,z)$ is given by

x	z	$p(x,z)$
0	0	0.15
0	1	0.35
1	0	0.25
1	1	0.25

marginalizing Z

marginalizing X

Further, we can compute

x	$p(x)$
0	0.5
1	0.5

and

z	$p(z)$
0	0.4
1	0.6

Conditional Distribution

Assume random variables $\underline{X} = \{X_1, X_2, \dots, X_D\}$ with joint distribution
$p_{\underline{X}}(x_1, x_2, \dots, x_D)$

Split $\{X_1, X_2, \dots, X_D\}$ into two arbitrary disjoint sets
$\underline{Y} = \{Y_1, Y_2, \dots, Y_I\}$ and $\underline{Z} = \{Z_1, Z_2, \dots, Z_J\}$, $\underline{Y} \cup \underline{Z} = \underline{X}$, $\underline{Y} \cap \underline{Z} = \{\}$.

The **conditional** distribution over $\underline{Y}$ given $\underline{Z}$ is given as

$$p_{\underline{Y}|\underline{Z}}(y_1, \dots, y_I \mid z_1, \dots, z_J) = \frac{p_{\underline{X}}(y_1, \dots, y_I, z_1, \dots, z_J)}{p_{\underline{Z}}(z_1, \dots, z_J)} \qquad \left(p_{\underline{Y}|\underline{Z}}(\underline{y} \mid \underline{z}) = \frac{p_{\underline{X}}(\underline{y}, \underline{z})}{p_{\underline{Z}}(\underline{z})} \right)$$

whenever $p_{\underline{Z}}(z_1, \dots, z_J) = p_{\underline{Z}}(\underline{z}) > 0$.

Alternatively, we can define conditional distributions as "first-class citizens", i.e. $p_{\underline{Y}|\underline{Z}}(\underline{y} \mid \underline{z})$ is a collection of distributions over $\underline{Y}$, one for each value $\underline{z}$ of $\underline{Z}$. Together with $p_{\underline{Z}}$ it gives rise to the joint $p_{\underline{X}}(\underline{y}, \underline{z}) := p_{\underline{Y}|\underline{Z}}(\underline{y} \mid \underline{z})\, p_{\underline{Z}}(\underline{z})$.

Example: Conditional Distribution

Recall the joint and marginal distributions from the previous example

x	y	z	p(x,y,z)
0	0	0	0.1
0	0	1	0.15
0	1	0	0.05
0	1	1	0.2
1	0	0	0.1
1	0	1	0.1
1	1	0	0.15
1	1	1	0.15

z	p(z)
0	0.4
1	0.6

$\div$

Then:

x	y	z	p(x,y|z)
0	0	0	0.25
0	1	0	0 125
1	0	0	0.25
1	1	0	0.375
0	0	1	0.25
0	1	1	0.333
1	0	1	0.166
1	1	1	0.25

Intuitive Semantics

Joint distribution $p_{\underline{X}}$

"Knowledge base," dependencies + uncertainty

Marginalization

"ignore," "forget," "account for"

Conditioning

"inject evidence," "observe"

(Conditional) Independence

Let $\underline{X}, \underline{Y}, \underline{Z}$ be random vectors (or single random variables).

$\underline{X}$ and $\underline{Y}$ are **independent** if the following equivalent* conditions hold:

- $p(\underline{x}, \underline{y}) = p(\underline{x}) \cdot p(\underline{y}) \quad \forall\, \underline{x}, \underline{y}$
- $p(\underline{x} \mid \underline{y}) = p(\underline{x}) \quad \forall\, \underline{x}, \underline{y}$
- $p(\underline{y} \mid \underline{x}) = p(\underline{y}) \quad \forall\, \underline{x}, \underline{y}$

observing one RV does not change our belief over the other RV

Moreover, $\underline{X}$ and $\underline{Y}$ are **conditionally independent** given $\underline{Z}$ if the following equivalent* conditions hold:

- $p(\underline{x}, \underline{y} \mid \underline{z}) = p(\underline{x} \mid \underline{z})\, p(\underline{y} \mid \underline{z}) \quad \forall\, \underline{x}, \underline{y}, \underline{z}$
- $p(\underline{x} \mid \underline{y}, \underline{z}) = p(\underline{x} \mid \underline{z}) \quad \forall\, \underline{x}, \underline{y}, \underline{z}$
- $p(\underline{y} \mid \underline{x}, \underline{z}) = p(\underline{y} \mid \underline{z}) \quad \forall\, \underline{x}, \underline{y}, \underline{z}$

* homework: show that the conditions are equivalent.

The Chain Rule

- Let $p(x_1, x_2, \ldots, x_D)$ be **any** distribution function.
- We can **always** factor it as

$$p(x_1, x_2, \ldots, x_D) = p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_2, x_1) \ldots p(x_D \mid x_1, \ldots, x_{D-1})$$

- In short $p(x_1, x_2, \ldots, x_D) = \prod_{i=1}^{D} p(x_i \mid x_1 \ldots x_{i-1})$

- This also works with **any ordering**, e.g.

$$\begin{aligned} p(x_1, x_2, x_3) &= p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_1, x_2) \\ &= p(x_2)\, p(x_3 \mid x_2)\, p(x_1 \mid x_2, x_3) \\ &= p(x_2)\, p(x_1 \mid x_2)\, p(x_3 \mid x_2, x_1) \\ &= \ldots \end{aligned}$$

Expectation

Let $\underline{X}$ be a random vector (or a single random variable).
The ***expectation*** (***expected value***) of $\underline{X}$ is defined as

$$\mathbb{E}[\underline{X}] = \sum_{\underline{x}} \underline{x}\, p(\underline{x}) \qquad \text{(discrete random variables)}$$

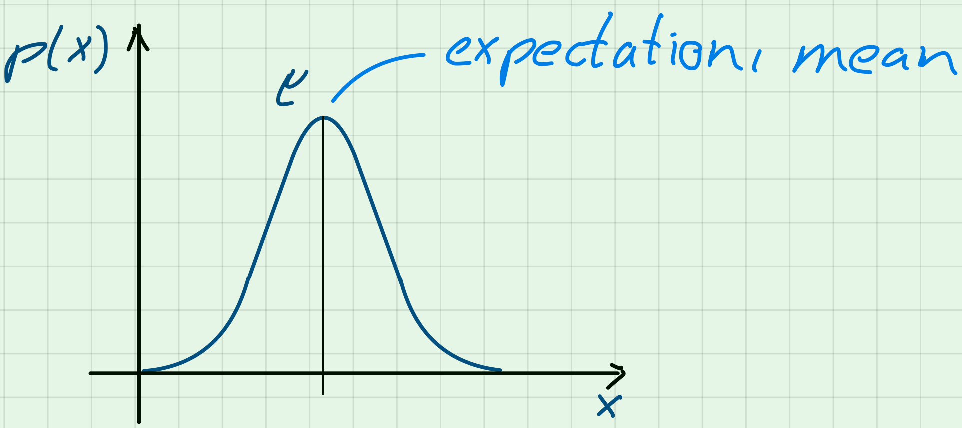
sum runs over all (joint) states of $\underline{X}$

$$\mathbb{E}[\underline{X}] = \int \underline{x}\, p(\underline{x})\, d\underline{x} \qquad \text{(continuous random variables)}$$

integral over whole state space of $\underline{X}$

Gaussian

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



Discrete Random Variable X

$\mathcal{X} = \{1, 2, 3\}$

$p(1) = 0.4\,, \quad p(2) = 0.3\,, \quad p(3) = 0.3$

$$\mathbb{E}[X] = 0.4 \times 1 + 0.3 \times 2 + 0.3 \times 3$$
$$= 1.9$$

Law of the Unconscious Statistician (LOTUS)

Let $\underline{X}$ be a random vector (or single RV) with distribution $p(\underline{x})$.
Let f be some function. Then

$$\underline{Y} := f(\underline{X})$$

is another random vector (or RV).

Then

often tricky to compute

$$\mathbb{E}[\underline{Y}] = \sum_{\underline{y}} p(\underline{y}) = \mathbb{E}[f(\underline{X})] = \sum_{\underline{x}} f(\underline{x})\, p(\underline{x}) \qquad \text{(discrete RVs)}$$

$$\mathbb{E}[\underline{Y}] \int p(\underline{y})\, d\underline{y} = \mathbb{E}[f(\underline{X})] = \int f(\underline{x})\, p(\underline{x})\, d\underline{x} \qquad \text{(continuous RVs)}$$

Often used as a definition, but actually a non-trivial result.

→ Law of the unconscious statistician

Conditional Expectation

Let $p(\underline{x} \mid \underline{y})$ be a conditional distribution over $\underline{X}$, conditional on $\underline{Y}$.

The **conditional expectation** $\mathbb{E}[\underline{X} \mid \underline{y}]$ is defined as

$$\mathbb{E}[\underline{X} \mid \underline{y}] := \int \underline{x}\, p(\underline{x} \mid \underline{y})\, d\underline{x} \qquad (\text{use } \Sigma \text{ for discrete RVs})$$

Of course, we can condition on many RVs as well:

$$\mathbb{E}[\underline{X} \mid \underline{y}, \underline{z}] = \int \underline{x}\, p(\underline{x} \mid \underline{y}, \underline{z})\, d\underline{x}$$

$$\mathbb{E}[\underline{X} \mid \underline{u}, \underline{v}, \underline{w}] = \int \underline{x}\, p(\underline{x} \mid \underline{u}, \underline{v}, \underline{w})\, d\underline{x}$$

$$\vdots$$

Iterated Expectation

Let $p(\underline{x} \mid \underline{y})$ be a conditional distribution over $\underline{X}$, conditional on $\underline{Y}$.
Let $p(\underline{y})$ be the (marginal) distribution over $\underline{Y}$.

Then

$$\mathbb{E}[\underline{X}] = \mathbb{E}_{\underline{Y}}[\mathbb{E}_{\underline{X}}[\underline{X} \mid \underline{Y}]] \qquad \text{(iterated expectation)}$$

(we use subscripts to make explicit which RVs we sum/integrate out)

<u>Proof:</u>

$$\underline{\mathbb{E}[\underline{X}]} = \int p(\underline{x})\, \underline{x}\, d\underline{x}$$

$$= \int \overbrace{\int p(\underline{x}, \underline{y})\, d\underline{y}}^{= p(\underline{x}) \text{ (marginal)}}\; \underline{x}\, d\underline{x}$$

$$= \int \int \overbrace{p(\underline{x} \mid \underline{y})\, p(\underline{y})}^{\text{chain rule}}\, d\underline{y}\; \underline{x}\, d\underline{x}$$

$$= \int \underbrace{\int p(\underline{x} \mid \underline{y})\, \underline{x}\, d\underline{x}}_{= \mathbb{E}_{\underline{X}}[\underline{X} \mid \underline{y}]} \cdot p(\underline{y})\, d\underline{y} = \underline{\mathbb{E}_{\underline{Y}}[\mathbb{E}_{\underline{X}}[\underline{X} \mid \underline{Y}]]}$$

Iterated Expectations cont'd

Iterated expectations generalize and can be "nested"

$$\mathbb{E}_{\underline{z}}\left[\mathbb{E}_{\underline{y}}\left[\mathbb{E}_{\underline{x}}[\underline{X} \mid \underline{Y}, \underline{Z}] \mid \underline{Z}\right]\right] = \int p(\underline{z}) \overbrace{\int p(\underline{y} \mid \underline{z}) \overbrace{\int \underline{x}\, p(\underline{x} \mid \underline{y}, \underline{z})\, d\underline{x}}^{= \mathbb{E}_{\underline{x}}[\underline{X} \mid \underline{Y}, \underline{Z}]}\, d\underline{y}}^{= \mathbb{E}_{\underline{Y}}[\mathbb{E}_{\underline{x}}[\underline{X} \mid \underline{Y}, \underline{Z}]]}\, d\underline{z}$$

$$= \int \underline{x} \underbrace{\iint \underbrace{p(\underline{x} \mid \underline{y}, \underline{z}) \cdot p(\underline{y} \mid \underline{z}) \cdot p(\underline{z})}_{p(\underline{x}, \underline{y}, \underline{z})}\, d\underline{y}\, d\underline{z}}_{p(\underline{x})}\, d\underline{x}$$

$$= \mathbb{E}[\underline{X}]$$

How does this simplify, when $\underline{X}$ is <u>conditionally independent</u> of $\underline{Z}$ given $\underline{Y}$?

Linear Algebra (extra slides)

Vector Space

A **vector space** is a set V on which addition $+$ is defined

- For any $\underline{u}, \underline{v} \in V$: $\underline{u} + \underline{v} \in V$
- For any $\underline{u}, \underline{v} \in V$: $\underline{u} + \underline{v} = \underline{v} + \underline{u}$ commutativity
- For any $\underline{u}, \underline{v}, \underline{w} \in V$: $\underline{u} + (\underline{v} + \underline{w}) = (\underline{u} + \underline{v}) + \underline{w}$ associativity

- There is a neutral element $\underline{0} \in V$, with $\underline{v} + \underline{0} = \underline{v}, \quad \underline{v} \in V$
- For each $\underline{v} \in V$, there is an inverse element $-\underline{v}$, with $\underline{v} + (-\underline{v}) = \underline{0}$

Moreover, multiplication with a scalar is defined

- For any $a, b \in \mathbb{R}$, $\underline{v} \in V$: $a(b\underline{v}) = (ab)\underline{v}$
- For any $\underline{u} \in V$: $1\underline{v} = \underline{v}$
- For any $a \in \mathbb{R}$, $\underline{u}, \underline{v} \in V$: $a.(\underline{u} + \underline{v}) = a\underline{u} + a\underline{v}$
- For any $a, b \in \mathbb{R}$, $\underline{v} \in V$: $(a+b)\underline{v} = a\underline{v} + b\underline{v}$

Canonical example: **Euclidean space** $\mathbb{R}^n$ $\quad \underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}$

Norm

Assigns a notion of length to vectors

Let V be a vector space. A **norm** on V is a function

$$\|\cdot\| : V \to \mathbb{R}$$

with

- $\|\underline{v}\| \geq 0 \qquad \underline{v} \in V$
- $\|\underline{v}\| = 0 \iff \underline{v} = \underline{0}$ — positive definiteness
- $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\| \qquad \underline{u}, \underline{v} \in V$ — triangle inequality
- $\|a \underline{v}\| = |a| \cdot \|\underline{v}\| \qquad a \in \mathbb{R}, \underline{v} \in V$ — homogeneity

For example, let $V = \mathbb{R}^N$.

Euclidean norm (ℓ_2-norm)

$$\|\underline{v}\|_2 = \sqrt{\sum_{k=1}^{N} v_k^2}$$

max-norm (union norm, infinity norm)

$$\|\underline{v}\|_\infty = \max_k |v_k|$$