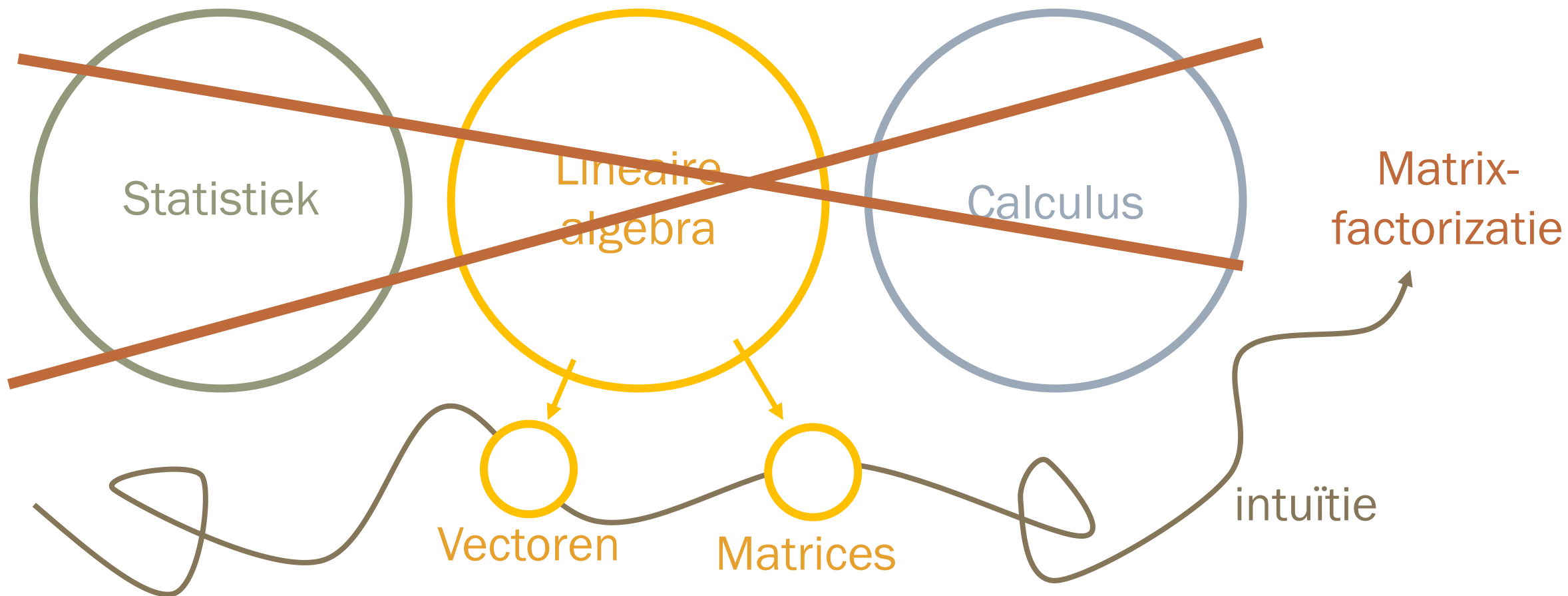




Collectieve Intelligentie

SIMON PAUW

MATRIXFACTORISATIE



Vandaag

Module 4:

- 1 Comedy films lijken op comedy films
 - One hot encoding
 - Jaccard
- 2 Films die lijken op films die je goed vind, vind je waarschijnlijk ook goed

Module 5:

- 1 Mensen die van comedy films houden, vinden een comedy film waarschijnlijk goed

Module 6:

- Gegeven dat we weten welke films je goed vindt, kunnen we bepalen van welke genres je houdt?

Uitgangspositie, vorige week:

	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

	drama	thriller	comedy
userId			
6	0.71	-0.65	0.56
7	0.85	0.22	0.18
8	0.89	-0.75	0.71
10	-0.02	0.20	0.02
11	0.94	0.33	0.07
12	0.94	0.18	0.29
13	0.75	0.75	0.29
14	0.07	0.04	0.49
15	0.52	0.42	-0.13
16	0.51	0.42	-0.31

userId	6	7	8	10	11	12	13	14	15	16
drama	0.71	0.85	0.89	-0.02	0.94	0.94	0.75	0.07	0.52	0.51
thriller	-0.65	0.22	-0.75	0.20	0.33	0.18	0.75	0.04	0.42	0.42
comedy	0.56	0.18	0.71	0.02	0.07	0.29	0.29	0.49	-0.13	-0.31

	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

	userId	6	7	8	10	11	12	13	14	15	16
movieName											
	Babe	0.62	0.75	0.80	0.04	0.80	0.86	0.84	0.27	0.42	0.34
	Inception	0.10	0.82	0.18	0.18	0.93	0.89	1.29	0.30	0.66	0.57
	A.I. Artificial Intelligence	0.30	0.71	0.41	0.08	0.80	0.78	0.91	0.17	0.51	0.47
	Ace Ventura: Pet Detective	0.45	0.25	0.58	0.05	0.17	0.35	0.45	0.48	-0.02	-0.19
	Bad Boys	0.67	0.84	0.87	0.09	0.86	0.97	1.07	0.45	0.44	0.30
	Changing Lanes	0.26	0.53	0.35	0.08	0.58	0.60	0.75	0.23	0.35	0.28
	Dumb & Dumber	0.45	0.22	0.58	0.05	0.13	0.32	0.42	0.50	-0.06	-0.24
	Event Horizon	-0.44	0.18	-0.50	0.15	0.25	0.16	0.58	0.07	0.30	0.29
	Full Metal Jacket	0.12	0.68	0.19	0.09	0.80	0.72	0.89	0.07	0.57	0.56
	I, Robot	-0.21	0.57	-0.20	0.20	0.67	0.60	1.11	0.24	0.55	0.48

```

1 user_matrix_transposed = user_matrix.T
2 predicted_ratings = movie_matrix @ user_matrix_transposed

```

Deze week

movieName			
Babe	?	?	?
Inception	?	?	?
A.I. Artificial Intelligence	?	?	?
Ace Ventura: Pet Detective	?	?	?
Bad Boys	?	?	?
Changing Lanes	?	?	?
Dumb & Dumber	?	?	?
Event Horizon	?	?	?
Full Metal Jacket	?	?	?
I, Robot	?	?	?

userId	6	7	8	10	11	12	13	14	15	16
	?	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?	?

Vaak hebben we geen genre informatie

Deze week

	userId	6	7	8	10	11	12	13	14	15	16
movieName											
Babe		0.80		1.00							
Inception										0.70	0.60
A.I. Artificial Intelligence			0.90							0.80	
Ace Ventura: Pet Detective		0.60							0.40		
Bad Boys		0.80									
Changing Lanes											
Dumb & Dumber		0.60		0.80					0.60		
Event Horizon								0.80			
Full Metal Jacket											0.90
I, Robot										0.70	

Maar wel *een deel* van de ratings

Stap 1

userId	6	7	8	10	11	12	13	14	15	16
	?	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?	?

movieName			
Babe	?	?	?
Inception	?	?	?
A.I. Artificial Intelligence	?	?	?
Ace Ventura: Pet Detective	?	?	?
Bad Boys	?	?	?
Changing Lanes	?	?	?
Dumb & Dumber	?	?	?
Event Horizon	?	?	?
Full Metal Jacket	?	?	?
I, Robot	?	?	?

movieName	userId	6	7	8	10	11	12	13	14	15	16
Babe		0.80		1.00							
Inception									0.70	0.60	
A.I. Artificial Intelligence			0.90						0.80		
Ace Ventura: Pet Detective		0.60						0.40			
Bad Boys		0.80									
Changing Lanes											
Dumb & Dumber		0.60		0.60				0.60			
Event Horizon							0.80				
Full Metal Jacket										0.90	
I, Robot									0.70		

$$M \cdot U^T = \hat{R}, \text{ gegeven } \hat{R} \text{ wat zijn } M \text{ en } U?$$

Step 1

userId	6	7	8	10	11	12	13	14	15	16
drama	0.71	0.85	0.89	-0.02	0.94	0.94	0.75	0.07	0.52	0.51
thriller	-0.65	0.22	-0.75	0.20	0.33	0.18	0.75	0.04	0.42	0.42
comedy	0.56	0.18	0.71	0.02	0.07	0.29	0.29	0.49	-0.13	-0.31

	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

	userId	6	7	8	10	11	12	13	14	15	16
movieName											
Babe		0.80		1.00							
Inception									0.70	0.60	
A.I. Artificial Intelligence			0.90						0.80		
Ace Ventura: Pet Detective		0.60						0.40			
Bad Boys		0.80									
Changing Lanes											
Dumb & Dumber		0.60		0.80				0.60			
Event Horizon							0.80				
Full Metal Jacket										0.90	
I, Robot									0.70		

Step 1

userId	6	7	8	10	11	12	13	14	15	16
drama	0.71	0.85	0.89	-0.02	0.94	0.94	0.75	0.07	0.52	0.51
thriller	-0.65	0.22	-0.75	0.20	0.33	0.18	0.75	0.04	0.42	0.42
comedy	0.56	0.18	0.71	0.02	0.07	0.29	0.29	0.49	-0.13	-0.31

	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

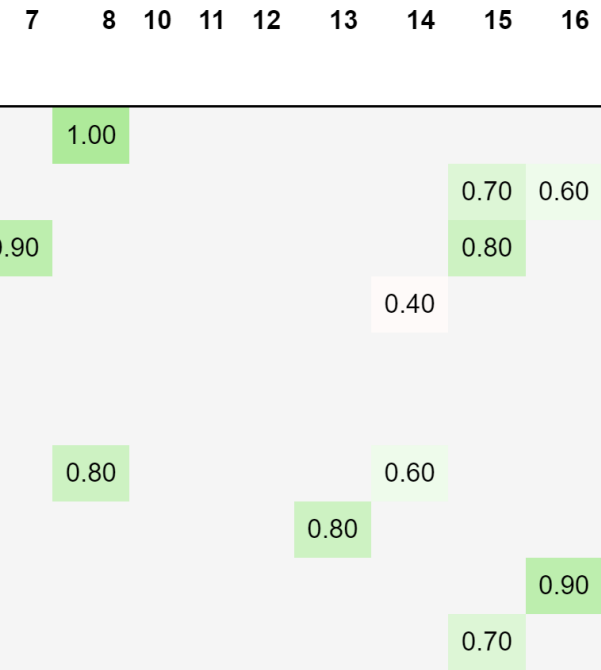
[illegible]

Stap 2

userId	6	7	8	10	11	12	13	14	15	16
drama	0.71	0.85	0.89	-0.02	0.94	0.94	0.75	0.07	0.52	0.51
thriller	-0.65	0.22	-0.75	0.20	0.33	0.18	0.75	0.04	0.42	0.42
comedy	0.56	0.18	0.71	0.02	0.07	0.29	0.29	0.49	-0.13	-0.31

	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

	userId	6	7	8	10	11	12	13	14	15	16
movieName											
	Babe	0.62	0.75	0.80	0.04	0.80	0.86	0.84	0.27	0.42	0.34
	Inception	0.10	0.82	0.18	0.18	0.93	0.89	1.29	0.30	0.66	0.57
	A.I. Artificial Intelligence	0.30	0.71	0.41	0.08	0.80	0.78	0.91	0.17	0.51	0.47
	Ace Ventura: Pet Detective	0.45	0.25	0.58	0.05	0.17	0.35	0.45	0.48	-0.02	-0.19
	Bad Boys	0.67	0.84	0.87	0.09	0.86	0.91	1.07	0.45	0.44	0.30
	Changing Lanes	0.26	0.53	0.35	0.08	0.58	0.60	0.75	0.23	0.35	0.28
	Dumb & Dumber	0.45	0.22	0.58	0.05	0.13	0.32	0.42	0.50	-0.06	-0.24
	Event Horizon	-0.44	0.18	-0.50	0.15	0.25	0.16	0.58	0.07	0.30	0.29
	Full Metal Jacket	0.12	0.68	0.19	0.09	0.80	0.72	0.89	0.07	0.57	0.56
	I, Robot	-0.21	0.57	-0.20	0.20	0.67	0.60	1.11	0.24	0.55	0.48



Stap 1
(matrixfactorisatie)

movieName	drama	thriller	comedy
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

userId	drama	thriller	comedy
6	0.71	-0.65	0.56
7	0.85	0.22	0.18
8	0.89	-0.75	0.71
10	-0.02	0.20	0.02
11	0.94	0.33	0.07
12	0.94	0.18	0.29
13	0.75	0.75	0.29
14	0.07	0.04	0.49
15	0.52	0.42	-0.13
16	0.51	0.42	-0.31


8

0

8

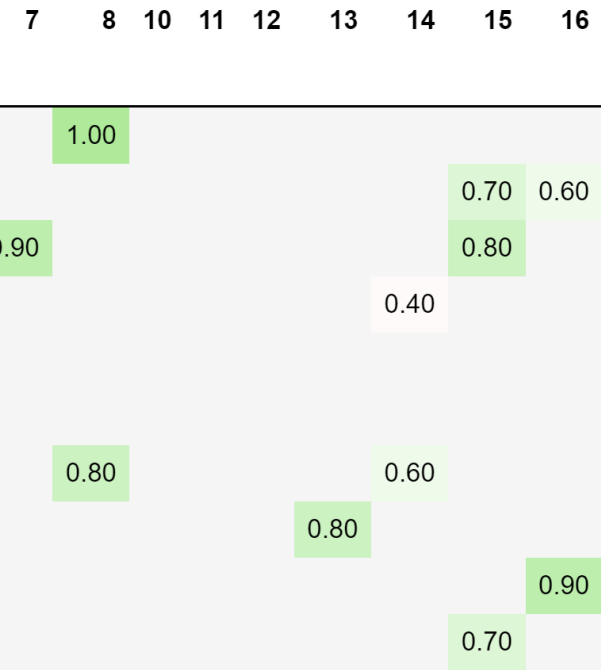
0

5



userId	6	7	8	10	11
movieName					
Babe	0.62	0.75	0.80	0.04	0.80
Inception	0.10	0.82	0.18	0.18	0.93
A.I. Artificial Intelligence	0.30	0.71	0.41	0.08	0.80
Ace Ventura: Pet Detective	0.45	0.25	0.58	0.05	0.17
Bad Boys	0.67	0.84	0.87	0.09	0.86
Changing Lanes	0.26	0.53	0.35	0.08	0.58
Dumb & Dumber	0.45	0.22	0.58	0.05	0.13
Event Horizon	-0.44	0.18	-0.50	0.15	0.25
Full Metal Jacket	0.12	0.68	0.19	0.09	0.80
I, Robot	-0.21	0.57	-0.20	0.20	0.67

Stap 2
(voorspellen)



Stap 1
(training)

movieName	drama	thriller	comedy
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

userId	drama	thriller	comedy
6	0.71	-0.65	0.56
7	0.85	0.22	0.18
8	0.89	-0.75	0.71
10	-0.02	0.20	0.02
11	0.94	0.33	0.07
12	0.94	0.18	0.29
13	0.75	0.75	0.29
14	0.07	0.04	0.49
15	0.52	0.42	-0.13
16	0.51	0.42	-0.31

Stap 2
(prediction)

	userId	6	7	8	10	11
movieName						
Babe	0.62	0.75	0.80	0.04	0.80	
Inception	0.10	0.82	0.18	0.18	0.93	
A.I. Artificial Intelligence	0.30	0.71	0.41	0.08	0.80	
Ace Ventura: Pet Detective	0.45	0.25	0.58	0.05	0.17	
Bad Boys	0.67	0.84	0.87	0.09	0.86	
Changing Lanes	0.26	0.53	0.35	0.08	0.58	
Dumb & Dumber	0.45	0.22	0.58	0.05	0.13	
Event Horizon	-0.44	0.18	-0.50	0.15	0.25	
Full Metal Jacket	0.12	0.68	0.19	0.09	0.80	
I, Robot	-0.21	0.57	-0.20	0.20	0.67	

Tussenstap

userId	6	7	8	10	11	12	13	14	15	16
	?	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?	?
	?	?	?	?	?	?	?	?	?	?

	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

	userId	6	7	8	10	11	12	13	14	15	16
movieName											
Babe	0.80		1.00								
Inception									0.70	0.60	
A.I. Artificial Intelligence		0.90							0.80		
Ace Ventura: Pet Detective	0.60							0.40			
Bad Boys	0.80										
Changing Lanes											
Dumb & Dumber	0.60		0.80					0.60			
Event Horizon							0.80				
Full Metal Jacket											0.90
I, Robot									0.70		

$$M \cdot U^T = \hat{R}, \text{ gegeven } \hat{R} \text{ en } M \text{ wat is } U?$$



Vraag 1: matrixfactorisatie

$$x \cdot b = y$$

$$x = 2 \text{ en } y = 6$$

wat is b ?



Matrixfactorisatie

$$X \cdot B = Y$$

matrices X en Y zijn bekend

wat is B ?

Vraag 2: dimensies

$$X \cdot B = Y$$

$$X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

wat zijn de dimensies van B ?



Matrixfactorisatie kan niet

$$X \cdot B = Y \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \longrightarrow \quad \text{Bevat geen oplossing!}$$

Wat zijn de dimensies van B? => wat zijn a en b?

Matrixfactorisatie kan niet

$$X \cdot B = Y \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \longrightarrow \quad \text{Bevat geen oplossing!}$$

Wat is B ?

$$X \cdot B = \hat{Y} \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Voor welk B ligt \hat{Y} zo dicht mogelijk bij Y ?

Matrixfactorisatie kan niet

$$X \cdot B = Y \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \longrightarrow \quad \text{Bevat geen oplossing!}$$

Wat is B ?

$$X \cdot B = \hat{Y} \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \longrightarrow \quad B \text{ zodat } \text{MSE}(\hat{Y}, Y) \text{ minimaal!}$$

Voor welk B ligt \hat{Y} zo dicht mogelijk bij Y ?

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waardes voor B .
- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waardes voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y}$$
$$B := B + \alpha \cdot X^T \cdot D$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + 0.01 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + 0.01 \cdot \begin{bmatrix} 1 \cdot 0 + 2 \cdot 4 + 3 \cdot -4 \\ 1 \cdot 0 + 3 \cdot 4 + 1 \cdot -4 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + 0.01 \cdot \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + \begin{bmatrix} -0.04 \\ 0.08 \end{bmatrix} = \begin{bmatrix} 5.96 \\ -1.92 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 5.96 \\ -1.92 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + \begin{bmatrix} -0.04 \\ 0.08 \end{bmatrix} = \begin{bmatrix} 5.96 \\ -1.92 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waardes voor B .

$$B = \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix} = \begin{bmatrix} 4.04 \\ 6.16 \\ 15.96 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waardes voor B .

$$B = \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix} = \begin{bmatrix} 4.04 \\ 6.16 \\ 15.96 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4.04)^2 + (10 - 6.16)^2 + (12 - 15.96)^2 = 30.43$$

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix}$$

- Herhaal:

- Bereken \hat{Y}
- Hoe fout is de oplossing?
- Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix} = \begin{bmatrix} 4.04 \\ 6.16 \\ 15.96 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4.04)^2 + (10 - 6.16)^2 + (12 - 15.96)^2 = 30.43$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4.04 \\ 6.16 \\ 15.96 \end{bmatrix} = \begin{bmatrix} -0.4 \\ 3.84 \\ -3.96 \end{bmatrix}$$

$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + 0.01 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} -0.4 \\ 3.84 \\ -3.96 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5.06 \\ -1.92 \end{bmatrix} = \begin{bmatrix} 4.04 \\ 6.16 \\ 15.96 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4.04)^2 + (10 - 6.16)^2 + (12 - 15.96)^2 = 30.43$$

$$D = Y - \hat{Y} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} - \begin{bmatrix} 4.04 \\ 6.16 \\ 15.96 \end{bmatrix} = \begin{bmatrix} -0.04 \\ 3.84 \\ -3.96 \end{bmatrix}$$

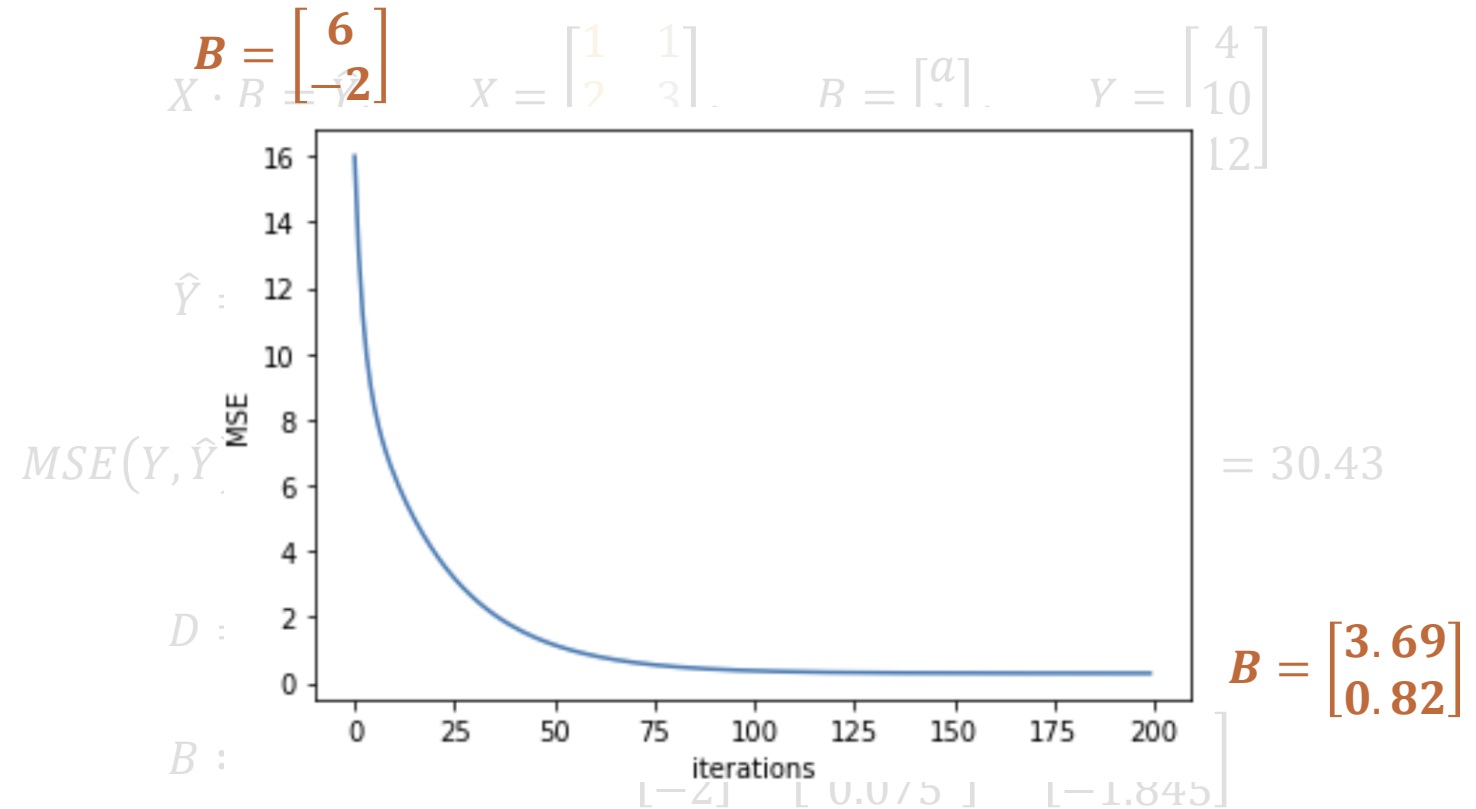
$$B := B + \alpha \cdot X^T \cdot D = \begin{bmatrix} 6 \\ -2 \end{bmatrix} + \begin{bmatrix} -0.042 \\ 0.075 \end{bmatrix} = \begin{bmatrix} 5.918 \\ -1.845 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .
- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.



Gradient descent

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waarden voor B .
- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3.69 \\ 0.82 \end{bmatrix} = \begin{bmatrix} 4.51 \\ 9.84 \\ 11.89 \end{bmatrix}$$

$$B = \begin{bmatrix} 3.69 \\ 0.82 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y}$$
$$B := B + \alpha \cdot X^T \cdot D$$

↑
Niet te groot!

Gradient descent

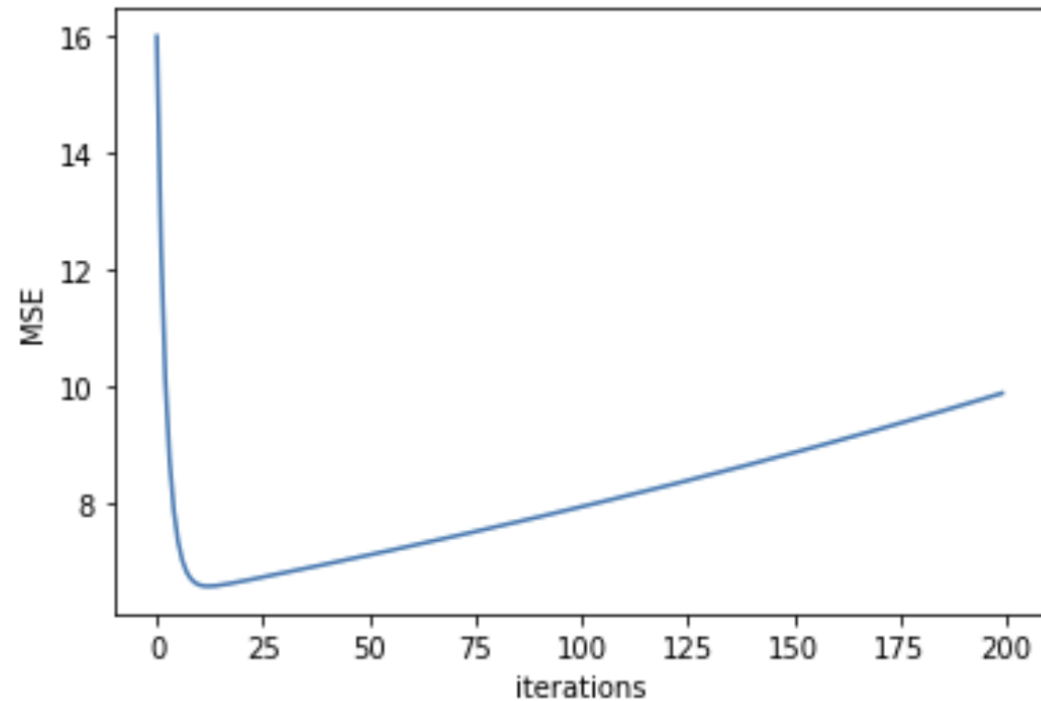
Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.



$$D = Y - \hat{Y}$$

$$B := B + \alpha \cdot X^T \cdot D$$

$$\Rightarrow \alpha = \cancel{0.01} 0.1$$

Gradient descent => movies

Welke B geeft de laagste MSE?

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Algoritme:

- Begin met random waarden voor B .
- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$\hat{Y} = X \cdot B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3.69 \\ 0.82 \end{bmatrix} = \begin{bmatrix} 4.51 \\ 9.84 \\ 11.89 \end{bmatrix}$$

$$\begin{array}{l} X \Rightarrow M \\ B \Rightarrow U^T \\ Y \Rightarrow R \\ \hat{Y} \Rightarrow \hat{R} \end{array}$$

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\begin{array}{l} X \Rightarrow M \\ B \Rightarrow U^T \\ Y \Rightarrow R \\ \hat{Y} \Rightarrow \hat{R} \end{array}$$

[illegible]

M	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

R	userId	6	7	8	10	11	12	13	14	15	16
	movieName										
	Babe	0.80		1.00							
	Inception								0.70	0.60	
	A.I. Artificial Intelligence		0.90						0.80		
	Ace Ventura: Pet Detective	0.60							0.40		
	Bad Boys	0.80									
	Changing Lanes										
	Dumb & Dumber	0.60		0.80					0.60		
	Event Horizon						0.80				
	Full Metal Jacket									0.90	
	I, Robot								0.70		

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}$$
$$B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\begin{aligned} X &\Rightarrow M \\ B &\Rightarrow U^T \\ Y &\Rightarrow R \\ \hat{Y} &\Rightarrow \hat{R} \end{aligned}$$

<i>M</i>	drama	thriller	comedy
movieName			
Babe	0.74	0.22	0.43
Inception	0.63	0.92	0.44
A.I. Artificial Intelligence	0.67	0.46	0.22
Ace Ventura: Pet Detective	0.05	0.18	0.95
Bad Boys	0.71	0.42	0.78
Changing Lanes	0.44	0.41	0.38
Dumb & Dumber	0.00	0.17	1.00
Event Horizon	0.00	0.74	0.08
Full Metal Jacket	0.66	0.53	0.00
I, Robot	0.34	1.00	0.35

<i>U^T</i>	userId	6	7	8	10	11	12	13	14	15	16
		?	?	?	?	?	?	?	?	?	?
		?	?	?	?	?	?	?	?	?	?
		?	?	?	?	?	?	?	?	?	?

<i>\hat{R}</i>	userId	6	7	8	10	11	12	13	14	15	16
movieName											
Babe	0.62	0.75	0.80	0.04	0.80	0.86	0.84	0.27	0.42	0.34	
Inception	0.10	0.82	0.18	0.18	0.93	0.89	1.29	0.30	0.66	0.57	
A.I. Artificial Intelligence	0.30	0.71	0.41	0.08	0.80	0.78	0.91	0.17	0.51	0.47	
Ace Ventura: Pet Detective	0.45	0.25	0.58	0.05	0.17	0.35	0.45	0.48	-0.02	-0.19	
Bad Boys	0.67	0.84	0.87	0.09	0.86	0.97	1.07	0.45	0.44	0.30	
Changing Lanes	0.26	0.53	0.35	0.08	0.58	0.60	0.75	0.23	0.35	0.28	
Dumb & Dumber	0.45	0.22	0.58	0.05	0.13	0.32	0.42	0.50	-0.06	-0.24	
Event Horizon	-0.44	0.18	-0.50	0.15	0.25	0.16	0.58	0.07	0.30	0.29	
Full Metal Jacket	0.12	0.68	0.19	0.09	0.80	0.72	0.89	0.07	0.57	0.56	
I, Robot	-0.21	0.57	-0.20	0.20	0.67	0.60	1.11	0.24	0.55	0.48	

$M \cdot U^T = \hat{R}$

Klaar, maar...

Geen exacte oplossing mogelijk! Waarom?

Waarom is dit de juiste update?

$$D = Y - \hat{Y}$$

$$B := B + \alpha \cdot X^T \cdot D$$

Waarom geen exacte oplossing?

Matrix factorisatie is niet hetzelfde als factorisatie van getallen

$X \cdot B = Y$ is niet hetzelfde als $x \cdot b = y$

Waarom geen exacte oplossing?

$$X \cdot B = Y \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \longrightarrow \quad \text{Bevat geen oplossing!}$$

Wat is B ?

$$X \cdot B = \hat{Y} \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \longrightarrow \quad B \text{ zodat } \text{MSE}(\hat{Y}, Y) \text{ minimaal!}$$

Voor welk B ligt \hat{Y} zo dicht mogelijk bij Y ?

Waarom geen exacte oplossing?

$$X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$X \cdot B = Y$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 1 \cdot a + 1 \cdot b \\ 2 \cdot a + 3 \cdot b \\ 3 \cdot a + 1 \cdot b \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

Vraag 3

$$1 \cdot a + 1 \cdot b = 4$$



Vraag 4

$$1 \cdot a + 1 \cdot b = 4$$

$$2 \cdot a + 3 \cdot b = 10$$



Vraag 5

$$1 \cdot a + 1 \cdot b = 4$$

$$2 \cdot a + 3 \cdot b = 10$$

$$3 \cdot a + 1 \cdot b = 12$$



Waarom geen exacte oplossing?

$$X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}, \quad X \cdot B = Y$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \cdot a + 1 \cdot b \\ 2 \cdot a + 3 \cdot b \\ 3 \cdot a + 1 \cdot b \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \end{bmatrix}, \quad X \cdot B = Y$$

$$\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \cdot a + 1 \cdot b \\ 2 \cdot a + 3 \cdot b \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y}$$
$$B := B + \boxed{\alpha \cdot X^T \cdot D}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

~~$$D = Y - \hat{Y}$$
$$B := B + \alpha X^T \cdot D$$~~

$$B = \begin{bmatrix} 6 + 0.001 \\ -2 \end{bmatrix} \uparrow$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 + 0.001 \\ 6 + 0.002 \\ 16 + 0.003 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4.001)^2 + (10 - 6.002)^2 + (12 - 16.003)^2 = 32.008$$

~~$$D = Y - \hat{Y}$$~~

~~$$B := B + \alpha \cdot X^T \cdot D$$~~

$$B = \begin{bmatrix} 6 + 0.001 \\ -2 \end{bmatrix} \Rightarrow MSE = 32.008$$



Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:

- Bereken \hat{Y}
- Hoe fout is de oplossing?
- Update: pas B een beetje aan.
 - Alternatief: probeer uit

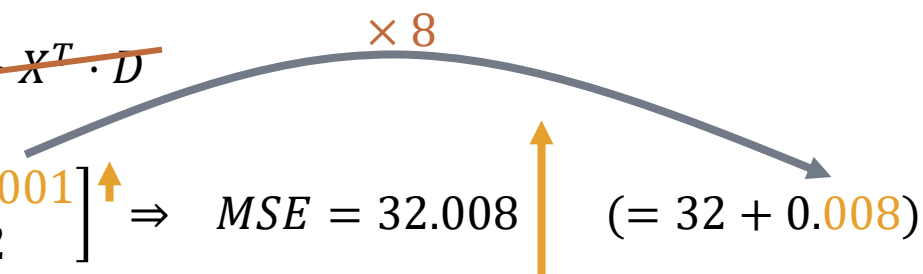
$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 + 0.001 \\ 6 + 0.002 \\ 16 + 0.003 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4.001)^2 + (10 - 6.002)^2 + (12 - 16.003)^2 = 32.008$$

~~$$D = Y - \hat{Y}$$~~

~~$$B := B + \alpha X^T \cdot D$$~~

$$B = \begin{bmatrix} 6 + 0.001 \\ -2 \end{bmatrix} \Rightarrow MSE = 32.008 \quad (= 32 + 0.008)$$


Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$


- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$\cancel{D = Y - \hat{Y}} \\ B := B + \cancel{\alpha \cdot X^T \cdot D}$$

$$B = \begin{bmatrix} 6 \\ -2 + 0.001 \end{bmatrix} \uparrow \Rightarrow MSE = 31.984 \quad (= 32 - 0.016)$$


Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$


- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$\cancel{D = Y - \hat{Y}} \\ B := B + \alpha \cancel{X^T \cdot D}$$

$$B = \begin{bmatrix} 6 \\ -2 + 0.001 \end{bmatrix} \uparrow \Rightarrow MSE = 31.984 \quad (= 32 - 0.016) \times -16$$


Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

~~$$D = Y - \hat{Y}$$~~

~~$$B := B + \alpha X^T \cdot D$$~~

Hoeveel invloed heeft een *kleine* verhoging van B op de MSE?

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix} \Rightarrow G = \begin{bmatrix} 8 \\ -16 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:

- Bereken \hat{Y}
- Hoe fout is de oplossing?
- Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

We willen de MSE zo klein mogelijk maken. Dus een klein stapje in de omgekeerde richting van G nemen.

~~$$D = Y - \hat{Y}$$~~

~~$$B := B + \alpha \cdot X^T \cdot D$$~~ $\Rightarrow B := B + \alpha \cdot -G$

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix} \Rightarrow G = \begin{bmatrix} 8 \\ -16 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y} \quad 2 \cdot X^T \cdot D \Leftrightarrow -G$$

$$B := B + \alpha \cdot X^T \cdot D \quad \Leftrightarrow \quad B := B + \alpha \cdot -G$$

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix} \Rightarrow G = \begin{bmatrix} 8 \\ -16 \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y}$$

$$B := B + \alpha \cdot X^T \cdot D \quad \Leftrightarrow \quad B := B + \alpha \cdot -G$$

$$X^T \cdot D = \begin{bmatrix} 1 \cdot (4 - 4) + 2 \cdot (10 - 6) + 3 \cdot (12 - 16) \\ 1 \cdot (4 - 4) + 3 \cdot (10 - 6) + 1 \cdot (12 - 16) \end{bmatrix}$$

Gradient descent

Welke B geeft de laagste MSE?

Algoritme:

- Begin met random waarden voor B .

$$B = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

- Herhaal:
 - Bereken \hat{Y}
 - Hoe fout is de oplossing?
 - Update: pas B een beetje aan.
 - Alternatief: probeer uit

$$X \cdot B = \hat{Y}, \quad X = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad Y = \begin{bmatrix} 4 \\ 10 \\ 12 \end{bmatrix}$$

$$\hat{Y} = \begin{bmatrix} 4 \\ 6 \\ 16 \end{bmatrix}$$

$$MSE(Y, \hat{Y}) = (4 - 4)^2 + (10 - 6)^2 + (12 - 16)^2 = 32$$

$$D = Y - \hat{Y}$$

$$B := B + \alpha \cdot X^T \cdot D \quad \Leftrightarrow \quad B := B + \alpha \cdot -G$$

$$X^T \cdot D = \begin{bmatrix} 1 \cdot (4 - 4) + 2 \cdot (10 - 6) + 3 \cdot (12 - 16) \\ 1 \cdot (4 - 4) + 3 \cdot (10 - 6) + 1 \cdot (12 - 16) \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

Vragen?

PROGIK@MPROG.NL