

Obsah

1	Vektory	3
1.1	Úvod do vektorov	3
1.1.1	Špeciálne príklady vektorov	4
1.1.2	Príklady využitia vektorov	4
1.2	Operácie s vektormi	5
1.2.1	Násobenie vektora skalárom	6
1.2.2	Lineárna kombinácia	7
1.2.3	Skalárny súčin vektorov	7
1.3	Normy vektora	10
2	Matice	15
2.1	Úvod do matíc	15
2.1.1	Príklady využitia matíc	16
2.1.2	Špeciálne príklady matíc	17
2.2	Operácie s maticami	18
2.2.1	Transponovanie matice	18
2.2.2	Sčítanie matíc	19
2.2.3	Násobenie matíc	19
2.2.4	Významné typy matíc	23
2.3	Systémy lineárnych rovníc	25
2.3.1	Riešenie pomocou rozšírenej matice	26
2.3.2	Elementárne riadkové operácie	26
2.3.3	Stupňovitý tvar matice	26
2.3.4	Lineárna závislosť	27
2.3.5	Počet riešení systému lineárnych rovníc	28
3	Vektorový priestor	29
3.1	Jadro a obraz matice	29
3.2	Báza	30
3.2.1	Ortogonálne vektory	30
3.3	Projekcia vektora	31
3.3.1	Gram-Schmidtov algoritmus	33
3.3.2	QR rozklad	34
3.4	Pseudoriešenia systému lineárnych rovníc	34
3.4.1	Pseudoriešenie pomocou ortogonálnej projekcie	34
3.4.2	Riešenie podurčeného systému lineárnych rovníc	35
3.4.3	Pseudoinverzia	35

4	Lineárne zobrazenia	37
4.1	Lineárne zobrazenia	37
4.1.1	Matica lineárneho zobrazenia	38
4.1.2	Zmena bázy	40
5	Vlastné vektory matice	43
5.1	Vlastný vektor	43
5.2	Diagonalizovateľné matice	45
5.2.1	Podobnosť matíc	45
5.2.2	Diagonalizovateľnosť matice	46
5.3	Singulárny rozklad	47
5.3.1	Aplikácie dekompozície na singulárne hodnoty	49
5.3.2	Riešenie sústav preurčených a podurčených rovníc	49
6	Strojové učenie a modelovanie	51
6.1	Strojové učenie	51
6.1.1	Základné pojmy	51
6.1.2	Metódy odhadu $f(X)$	52
6.1.3	Pravdepodobnostný pohľad na funkciu f	52
6.1.4	Základy klasifikácie	52
6.1.5	Metóda K najbližších susedov	53
6.2	Kompromisy v modelovaní	55
6.2.1	Kompromis medzi vychýlením a rozptylom	56
6.2.2	Rozšírenie o validačnú množinu	58
6.2.3	Krížová validácia	58
6.2.4	Stratifikované vzorkovanie	60
7	Lineárna regresia	61
7.1	Jednoduchá lineárna regresia	61
7.2	Viacnásobná lineárna regresia	62
7.2.1	Odhad koeficientov viacnásobnej lineárnej regresie	63
7.2.2	Vyhodnotenie odhadu	64
7.2.3	Klasifikácia pomocou lineárnej regresie	64
7.2.4	Odhady regresných koeficientov pri neexistujúcej inverzii a numericky nestabilných výpočtoch	64
7.3	Metóda gradientového poklesu	65
7.3.1	Princíp metódy gradientového poklesu	65
7.3.2	Metóda gradientového poklesu pre lineárnu regresiu	67
8	Príprava dát na strojové učenie	68
8.1	Inžinierstvo príznakov	68
8.1.1	Neurálne embeddingy	69
8.1.2	Embeddingy na škálovanie	70
8.1.3	Vektorové embeddingy	70
8.2	Inžinierstvo a transformácie príznakov	72
8.2.1	Transformácie príznakov	72
8.3	Redukcia dimenzionality	75
8.3.1	Metóda hlavných komponentov	75
8.3.2	Doplňanie chýbajúcich hodnôt pomocou metódy hlavných komponentov	78

9	Vyhodnocovanie úspešnosti modelov	80
9.1	Miery úspešnosti pre regresiu	80
9.1.1	Miery zohľadňujúce počet príznakov	81
9.2	Vyhodnocovanie úspešnosti pre binárnu klasifikáciu	82
9.2.1	Matica zámen	82
9.2.2	Miery úspešnosti binárnej klasifikácie	82
9.2.3	Komplexnejšie miery presnosti binárnej klasifikácie	84
9.2.4	ROC analýza	85
9.2.5	Rozšírenie na viactriednu klasifikáciu	85
10	Minimalizácia empirického rizika	87
10.1	Stratové funkcie pre regresné úlohy	88
10.1.1	Robustné stratové funkcie	89
10.1.2	Stratové funkcie pre binárnu klasifikáciu	91
11	Regularizácia	96
11.1	Regularizovaná minimalizácia empirického rizika	96
11.2	Populárne regularizované metódy	97
11.2.1	Hrebeňová regresia	97
11.2.2	Lasso	99
11.2.3	Elastic Net regresia	102
12	Metódy nelineárnych modelov	104
12.1	Metódy založené na rozhodovacích stromoch	104
12.1.1	Regresné stromy	104
12.1.2	Bagging	107
12.1.3	Náhodný les	109
12.1.4	Boosting	109
12.2	Neurónové siete	111
12.2.1	Dopredné neurónové siete	111
12.2.2	Klasifikácia pomocou neurónových sietí	112
13	Učenie bez učiteľa	115
13.1	Funkcie nehodnovernosti	115
13.1.1	Vybrané aplikácie pre l^q a l^a	115
13.2	K-Means	116
13.2.1	Výber počtu zhlukov – metóda lakťa	117
13.2.2	Vybrané aplikácie pre K-means	118
13.3	Hierarchické zhľukovanie	119
13.4	DBSCAN – Zhľukovanie založené na hustote	120