

Miniprojekt z Regresji Liniowej

Maciej Brzozowski

8 kwietnia 2024

1 Wprowadzenie

- Dane zostały przeskalowane metodą min-max 1 (standaryzacja nie wykazywała znaczących różnic w wynikach).
- Użyta została strata kwadratowa.
- Do rozwiązania analitycznego dane zostały podzielone na zbiór treningowy i testowy w stosunku 80/20.
- Do metod gradientowych dane zostały podzielone na zbiór treningowy, walidacyjny i testowy w stosunku 50/25/25
- W raporcie przedstawione zostały wyniki z następujących funkcji:
 - X
 - $\sum_i^5 X^i$
 - $\sum_i^{10} X^i$
 - $X \times X$
 - $\sum_i^{10} (X \times X)^i$
 - $\sum_i^5 X^i \times \sum_i^5 X^i$
 - $\exp(-(X - \bar{X}))$

2 Rozwiązania analityczne

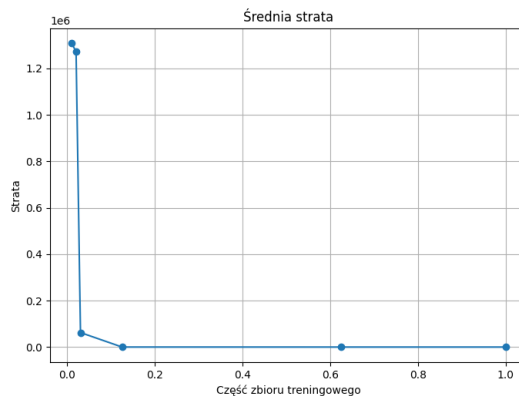
Poniższa tabela przedstawia wartości funkcji straty kwadratowej dla różnych funkcji bazowych dla danych przeskalowanych przez min-max 1:

$f(X)$	Błąd na zbiorze treningowym	Błąd na zbiorze testowym	czas uczenia
X	128814	129936	0.5s
$\sum_i^5 X^i$	13446	13606	0.5s
$\sum_i^{10} X^i$	235	226	0.5s
$X \times X$	123243	128912	0.5s
$\sum_i^{10} (X \times X)^i$	35	40	4.4s
$\sum_i^5 X^i \times \sum_i^5 X^i$	24	44	35.2s
$\exp(-(X - \bar{X}))$	129671	129819	0.5s

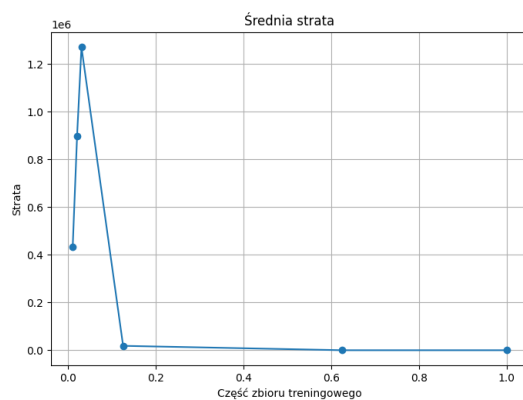
Dane zostały podzielone na zbiór treningowy i testowy w stosunku 80/20. Do odwrócenia macierzy została użyta odwrotność Moore-Penrose. Próby wykorzystania regularyzacji l_2 dla najlepszych funkcji bazowych nie wykazały poprawy wyników.

2.1 krzywe uczenia

Poniższy wykres przedstawia krzywą uczenia dla funkcji $\sum_i^{10} X^i$



Poniższy wykres przedstawia krzywą uczenia dla funkcji $\sum_i^{10} (X \times X)^i$



W okolicach $\frac{1}{8}$ zbioru treningowego strata modelu drastycznie maleje

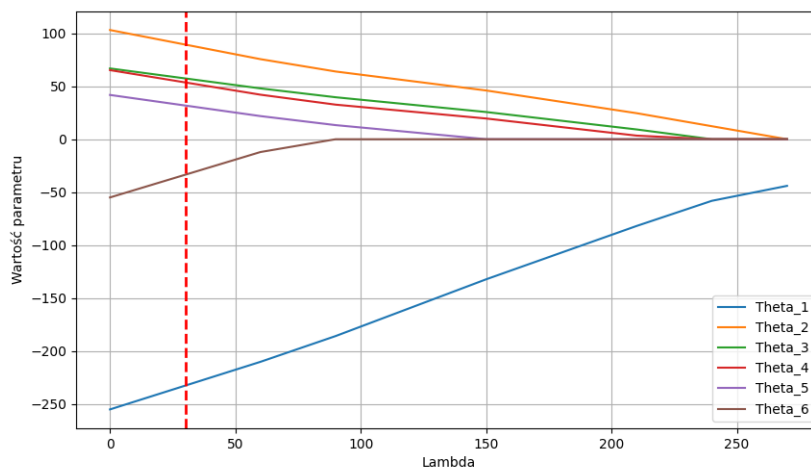
3 Metoda gradientowa

Poniższa tabela przedstawia wartości funkcji straty kwadratowej dla różnych funkcji bazowych dla danych przeskalowanych przez min-max 1:

$f(X)$	Błąd na zbiorze treningowym	Błąd na zbiorze walidacyjnym	czas uczenia
X	135158	124653	0.7s
$\sum_i^5 X^i$	107161	102173	4.1s
$\sum_i^{10} X^i$	74135	62968	25.9s
$X \times X$	128946	122568	6.0s
$\sum_i^{10} (X \times X)^i$	98177	98367	22.8s
$\sum_i^5 X^i \times \sum_i^5 X^i$	102660	108216	60.5s
$\exp(-(X - \bar{X}))$	135770	125724	1.5s

3.1 Regularyzacja

Z racji tego, że błąd uzyskany bez regularyzacji na zbiorze treningowym nie bardzo różni się od błędu na zbiorze walidacyjnym, zdecydowałem się skupić na regularyzacji lasso w celu redukcji parametrów. Poniższy wykres przedstawia ścieżki regularyzacji dla funkcji identycznościowej.



4 Podsumowanie

- Uzyskane wyniki są o wiele lepsze dla rozwiązania analitycznego.
- Funkcje bazowe o wysokim stopniu wielomianu i zawierające iloczyn kartezjański drastycznie zmniejszają stratę
- Regularyzacja nie wpływa znacząco na wyniki
- Dla rozwiązań analitycznych $\frac{1}{8}$ zbioru treningowego (ok. 200 rekordów) wystarcza aby uzyskać sensowny model