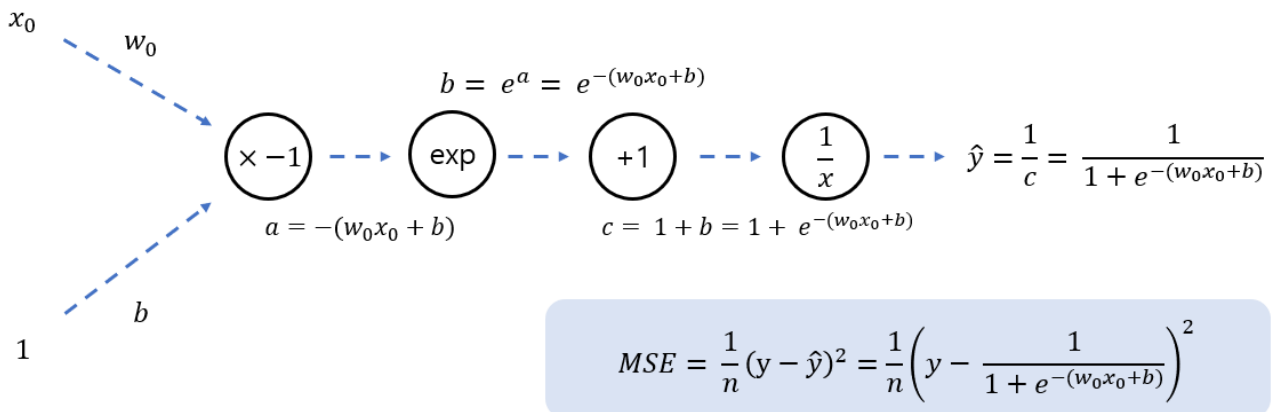


- 이번 주차 과제는 이론의 깊은 이해를 위해 직접 계산해보는 문제가 많습니다. 문제를 풀면서 문제가 의미하고 있는 바에 대해 집중하면서 풀어주시면 더 좋을 것 같습니다. 코드 작성 문제의 경우 R을 이용해주시면 됩니다.
- 제출형식은 PDF 및 html 파일로 변환해주시면 감사하겠습니다. 파일은 psat2009@naver.com으로 보내주세요.
- 제출기한은 2월 22일 **목요일 자정까지** 입니다. 패키지와 마찬가지로 무단 미제출 2회 시 퇴출이니 유의해주세요.

Chapter 1 : 역전파(BackPropagation)

2주차 전반부 리드오프에서는 Gradient Descent 알고리즘이 응용되는 분야인 딥러닝의 역전파에 대해 알아보았습니다. 역전파 알고리즘의 이해는 딥러닝의 원리, 그리고 추후 나올 여러 활성화함수를 이해하기 위해서도 매우 중요합니다. 따라서 비교적 간단한 예시와 함께, R로 직접 알고리즘을 구현해보는 연습을 해보겠습니다. 그림의 a, b, c 는 Chain Rule 적용 등 문제 풀이에 도움이 될 수 있도록 임의로 넣은 파라미터입니다. 실제로 그래디언트를 계산하고 업데이트할 파라미터는 w_0, b 이니 참고해주세요.

로지스틱 회귀의 연산 과정은 다음과 같이 도식화할 수 있습니다. 여기서 w_0, b 가 우리가 조정하고자 하는 파라미터가 될 것입니다. 초기 w_0, b 를 각각 -1, 2로, 인풋 $x_0 = -1$, 실제 $y = 1$ 이라고 하겠습니다.



문제1. 인풋 $x_0 = -1$ 이라고 했을 때, a, b, c, \hat{y} 은 각각 얼마인지 계산해주세요. Loss는 얼마인지 계산해주세요.

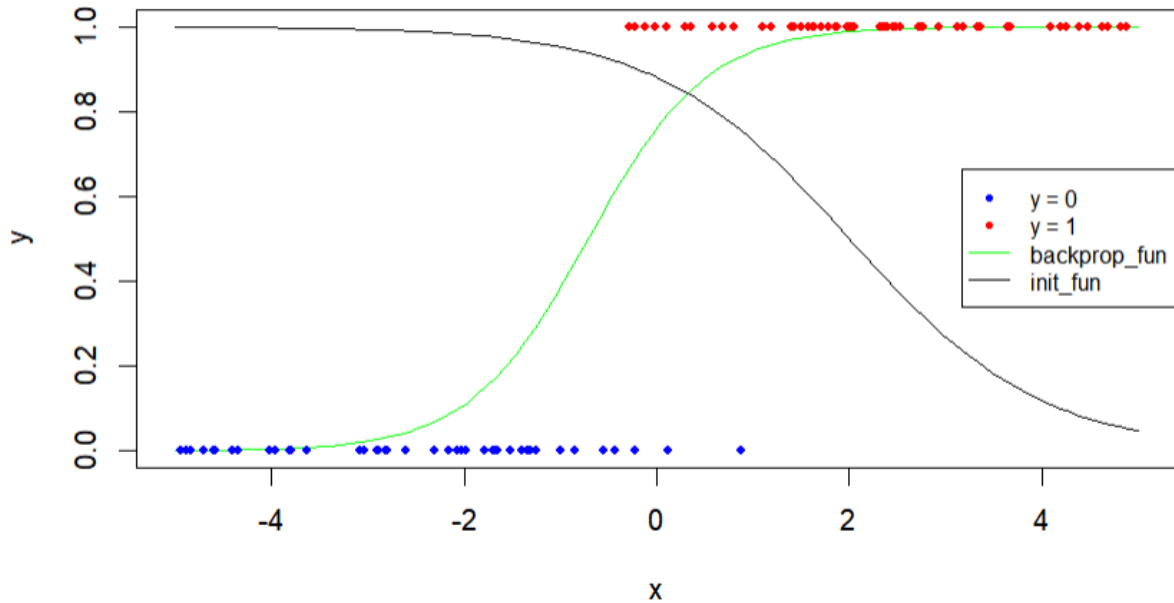
문제2. $\frac{\partial L}{\partial w_0}$, $\frac{\partial L}{\partial b}$ 를 각각 계산해주세요. 반드시 $\frac{\partial L}{\partial \hat{y}}$, $\frac{\partial \hat{y}}{\partial c}$ 등을 이용하여 Chain Rule을 사용해주세요.

문제3. 1주차의 Gradient Descent 알고리즘을 remind 해봅시다. learning rate를 0.1로 했을 때, 역전파가 한 번 수행된 후 업데이트된 w'_0, b' 은 각각 어떻게 될지 계산해주세요.

(보너스문제1). 역전파를 통해 weight와 bias를 수정한 뒤, 이진 분류 문제를 수행하고자 합니다. 아래의 링크에 접속해주세요. 링크에 있는 코드는 역전파 알고리즘을 수행하는 함수입니다. 코드의 빈칸을 채워주세요.

<https://maize-shield-367.notion.site/2-Chapter1-f56caa659ab24351a470e8620cd07364?pvs=4>

(보너스문제2). 역전파를 통해 fitting한 시그모이드 함수, fitting 이전의 함수, y값들을 다음과 같이 plot에 나타내주세요. (형식은 자유입니다)



Chapter 2: 라그랑주 승수법

2주차 전반부에 라그랑주 승수법에 대해서도 가볍게 살펴보았습니다. 이는 제약이 있는 문제를 제약이 없는 문제로 바꾸어 보다 문제를 쉽게 해결할 수 있도록 한다는 점에서 의미가 있었습니다. 다소 어려운 개념이었기 때문에, 라그랑주 승수법을 직접 구현하기보다 해당 기법이 사용된 예시를 통해 복습해보도록 하겠습니다.

문제1. Ridge와 달리, LASSO는 회귀계수의 절댓값 합에 제약을 거는 방법입니다. 즉 제약식이 $\sum_{j=1}^p |\beta_j| \leq s$ 임을 고려했을 때, LASSO의 목적함수는 어떻게 될지 생각해본 후, 작성해주세요.

문제2. Ridge Regression을 R로 수행해보는 코드를 작성해보겠습니다.

- 'glmnet' 패키지를 불러와주세요.
- Carseats 데이터를 data에 저장해주세요.
- seed는 3233으로 고정해주세요.

문제 2-1. Sales 변수에 대해 회귀모델을 적합하고자 합니다. 데이터셋에서 Factor형 변수를 제거해주세요. 총 3개의 열이 제거됩니다.

문제 2-2. Sales 변수를 y로 두어, 일반 회귀 모델을 적합해주세요. lm() 함수를 사용하면 됩니다.

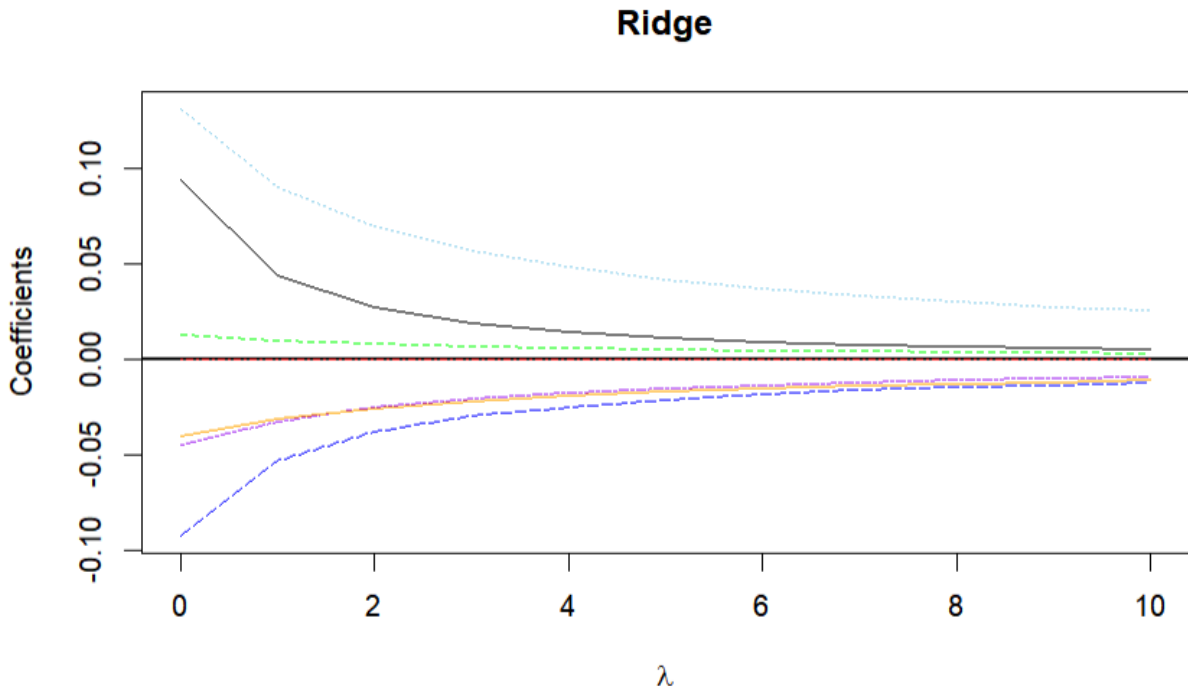
문제 2-3. glmnet() 함수를 사용하여, Ridge 회귀 모델을 적합해주세요. alpha = 0, lambda = 3으로 설정해주세요.

문제 2-4. 선형 회귀와 Ridge 계수를 비교할 수 있는 matrix 또는 데이터프레임을 만들어 보여주세요.

(HINT) coef() 함수와 cbind()를 활용하면 쉽게 비교할 수 있습니다.

문제 2-5. $\lambda=0$ 으로 설정하여 Ridge Regression을 수행해보고, 계수가 어떻게 변하는지 파악해주세요.

(보너스문제3). λ 를 0부터 10까지 1만큼 증가시키면서, 회귀계수들이 어떻게 변하는지 시각화해주세요. (형식은 자유입니다)



(HINT) `lambda_seq`의 경우, `seq()` 함수를 사용하면 쉽게 선언할 수 있습니다.

(HINT) 반복문을 통해 `lines()`를 그리면 보다 간단하게 plot을 그릴 수 있습니다.

Chapter 3: 행렬과 선형 변환

2주차 후반부 리드오프를 통해, 선형 변환과 기저벡터 변화의 관점에서 행렬을 이해하면 여러 가지 행렬의 기하학적 의미에 대해서도 이해가 가능해진다는 것을 배웠습니다. 이러한 관점에서 선형대수학을 바라보는 것은 추후에 나올 고유값과 고유값 분해, 그 응용인 PCA의 이해를 위해서도 매우 중요합니다. 따라서 Chapter3은 R로 코드를 구현하기보단, 좀 더 깊은 이해를 위해 간단한 문제를 풀어보는 시간을 갖도록 하겠습니다.

문제1. 공분산행렬 $X^T X$ 가 $\begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix}$ 과 같이 주어졌습니다. Gaussian Distribution을 통해 랜덤하게 분포된 원형의 데이터는 위 공분산벡터에 의해 어떻게 변환될지 상상하여 간단하게 그려주세요. (계산이 필요없습니다)

문제2. 2차원 데이터(벡터)들을 반시계방향으로 $\frac{\pi}{2}$ 만큼 회전시키는 2×2 행렬(A)을 작성해주세요. 또 $x=y$ 직선에 대해 대칭시키는 2×2 행렬(B)을 작성해주세요. 벡터에 대해 A를 먼저 적용하고 B를 적용하는 것은 B를 먼저 적용하고 A를 적용하는 것과 왜 다른지($BAx \neq ABx$) 표준기저벡터 변화의 관점에서 간단하게 설명해주세요. (계산이 필요없습니다)

(HINT) $x=y$ 직선에 대해 대칭시킨다는 것은 x, y 좌표가 서로 바뀐다는 것을 의미합니다.)

(보너스문제4). 3차원 계에서, z 축은 고정시키고 데이터(벡터)들을 xy 평면에서 반시계방향으로 θ 만큼 회전시키는 3×3 행렬을 작성해주세요.

문제3. 대각행렬의 주대각성분의 곱은 대각행렬의 행렬식과 같음을 표준기저벡터 변화의 관점에서 간단하게 설명해주세요. 2×2 행렬에 대해서만 보여주셔도 좋습니다. (행렬식 계산 공식 이용 X)

문제4. 투영 행렬 $H (X(X^T X)^{-1} X^T)$ 가 idempotent($H^2 = H$) 하다는 것을 계산해주세요. 투영 행렬이 y 를 X 의 열공간에 투영시키는 변환임을 고려했을 때, H 가 idempotent하다는 것은 기하학적으로 무엇을 의미하는지 생각해본 후 간단하게 작성해주세요.

(보너스문제5). 선형회귀분석을 수행하고자 합니다. X 의 Design Matrix가 정방행렬이며 또 full rank matrix 라고 가정해봅시다. 예를 들어 $\begin{bmatrix} a & 1 \\ b & 1 \end{bmatrix}$ ($a \neq b$) 와 같은 행렬 X 에 대해, 투영행렬 $H(X(X^T X)^{-1} X^T)$ 가 I (identity matrix)와 같음을 보여주세요.

Design Matrix X 의 열의 수가 추정하고자 하는 모수(β) 개수, 행의 수가 데이터 개수임을 고려했을 때, 위 결과가 의미하는 바가 무엇인지 생각해본 후 간단하게 설명해주세요.