경사 하강법이란 어떠한 현상을 가장 잘 설명 할 수 있는 어떠한 수식을 찾아내고자 할 때 그래프의 성분을 편미분과 차이의 제곱 이러한 개념들을 사용하여 원하는 그래프의 근사해로 수렴시키는 방법을 말한다.

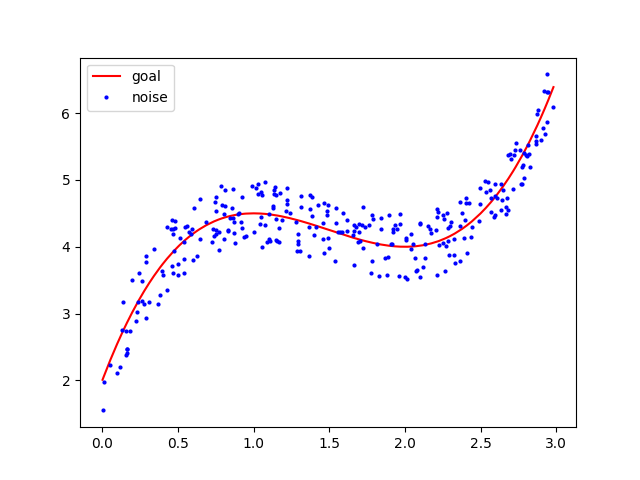
이번 과제에서는 y=x3-4.5x2+6x+2 라는 정답 값에 노이즈를 의도적으로 유발하고 그 노이즈 만을 이용하여 원래 정답 값을 찾아가도록 하는 실험이다.

Stepsize 는 0.001

초기 gradient는 [1,5]중 랜덤을 사용하여 임의로 지정하였다.

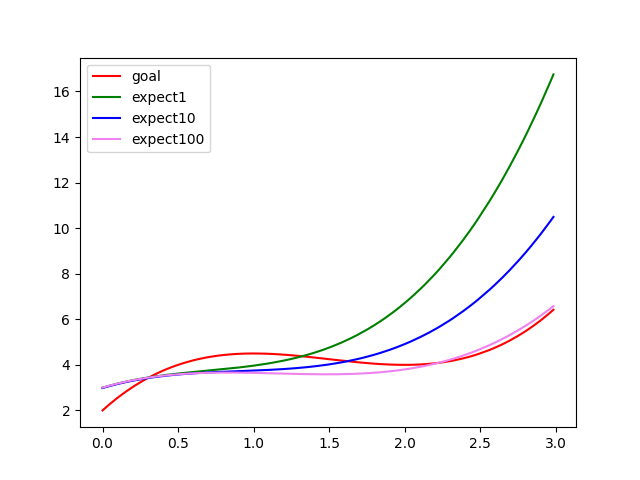
아래는 각각

정답 값과 [-0.5, 0.5]구간으로 발생시킨 잡음이다.



이러한 환경에서 goal 값에 수렴시키기 위하여 각각 1,10,100번 반복하여 최적화를 진행 한 결과이다.

많은 반복을 거칠수록 정답 값에 가까워 진다.



이번과제에서는 이전과는 달리 반복수를 많이 가져가는 단계에서 스텝 사이즈를 줄일 필요가 없다고 느껴졌다. 반복을 거칠수록 정답과의 차이가 작아지고 그렇게 되면 계산되는 기울기가 작아지며 알아서 미세한 조정을 하게 된다.

그러나 반복을 초월적으로 많이 가져가게 되면서 알아차린 사실은 cost값이 감소만 하는 것이 아니라는 것이었다.

기울기의 반대 방향으로 움직이게 되니 반드시 감소하리라 생각되었지만 무분별하게 반복수를 늘려나가면 다시 cost값이 커지며 그래프는 발산하게 된다.

그리하여 마지막 실험을 기획 하게 되었다.

마지막 실험이 이전과 다른 것은 조건에 따른 무한 반복문이라는 것에 있다. Cost가 이전과 비교하여 10번이상 커진다면 스텝 사이즈를 줄인다. 스텝사이즈가 크다면 기울기가 수렴하는지점을 넘어가는 경우가 발생한다. 반면에 스텝사이즈가 작다면 생기는 문제는 반복을 더 많이 가져가면 해결 된다.

만약 이전과 cost가 같다면 반복을 종료한다 이것은 더 이상 개선되지 않는 수렴이라고 판단한다.

그리고 함수의 차수를 두 차원 더 늘려 5차함수 꼴로 그래프를 정의한다. 특징 벡터가 많을수록 더욱 현상을 잘 설명할수 있기 때문이다.

이러한 조건에 따라 충분한 시간이 주어진다면 반드시 수렴하는 궁극의 시스템이 완성된다.

충분히 작지 않은 값에서 수렴이라 판단될 경우 스텝 사이즈를 키워 지역해를 탈출 시키는 방안도 고려되었지만 오히려 학습을 원점으로 되돌리는 현상이 되어 폐기하였다.

스텝 사이즈를 0.00001 이하로 내려가지 않게 한계를 두는 전략을 고안해 보았지만 y와 yhat값이 비슷해짐에 따라 기울기 값이 너무나 작은 수가 되어 한없이 작은수를 하염없이 조정하는 시간을 보내게 된다. 그리하여 영상에는 담지 않았다.

아래 링크는 수렴을 보이는 영상(직접 녹화 하였음)의 링크이며 빠르고 안정적으로 수렴된다.

https://youtu.be/3BZeOaiyYOg?si=Oiw1nW7bl8vMgUbV