

미적분 수행평가 보고서

31107 김유빈

- 주제 : 인공지능의 발전으로 인한 전력사용량 증가를 수학적으로 예측하는 방법

<요약>

미래의 ai로 인해 증가할 전력 사용량 추이를 구해야 함 > 주어진 데이터에 맞는 가장 적절한 추세를 예측하는 기법 : 선형회귀 > 미래의 추세를 예측할 수 있음. > 그럼 가장 적절한 추세선을 어떻게 그음? > 가장 적절한 선을 긋는 방법 - 최적화(그 중 경사하강법) > 평년의 데이터 분석 변수가 1개일 때는 이계도함수가 0이 될 때에서 가장 안정한 상태가 되는 것을 설명할 수 있다. > 1번 2000 ~ 2030 전력 사용량 추이 및 예측 그래프 제시 > 하지만, ai로 인해 전력 사용량이 크게 변화한다는 것을 체감하기 힘들 > because ai로 인해 우리가 구해야하는 전력 사용량 추이는 다양한 변수가 있음 > 변수가 2개 이상이 되는 2차원 이상의 데이터에서는 일반적인 미분법으로는 계산을 할 수 없기 때문에 편미분을 도입해야 한다. 편미분을 도입해야 하는 이유 - 쌍곡포물선의 예시 (x축의 관점으로 보면 극대, y축의 관점으로 보면 극소) >>우리가 배운 음함수, 매개변수 미분법으로는 설명할 수 없다. > 그럼 어떤 미분법이 필요한가? > 편미분 > 편미분에 대한 간단한 예시 > 편미분을 사용한 2,3차원 미분 > 그래디언트 벡터의 방향을 구할 수 있음 > 하지만 우리가 구해야 하는 전력 사용량 추이는 많은 데이터를 필요로 함 > 계산량이 너무 많아짐 > 해결방안 : 전체데이터를 미분하는 것이 아니라, 데이터의 일부만 미분해도 되지 않을까? > 확률적 경사하강법 > 이를 통해 구한 ai 로 인해 증가량 전력 사용량 추이 (직접 계산하려 했음 but > 신뢰하기 힘들고, 많은 데이터를 처리할 수 있어야 함 > 따라서 믿을 수 있는 기관의 데이터 제시 1. 데이터 센터에서 소비할 데이터 추이 2030년에 소비할 것으로 예측되는 전력량 1200 테라와트 = 2021년 기준 한국이 2년동안 사용한 전력량과 같다. 2. 2030년까지 전세계 전력 소비 증가율 중 데이터 센터가 차지할 비율 매년 2.4% 중 약 40% 차지함) 따라서 우리는 인공지능을 발전시키는데만 집중하는 것이 아닌, 인공지능이 사용할 전기를 어떻게 친환경적으로 발전할 것인지에 대해서도 많은 논의와 발전이 필요하다.

<주제 선택 이유>

최근 뉴스와 유튜브를 통해 인공지능의 급속한 발전과 그로 인한 데이터 센터 수요의 증가에 대한 소식을 접하게 되었습니다. 이러한 증가로 인해 전력 부족 현상이 심화되고 있으며, 이는 환경 파괴를 가속화하고 있습니다. 이와 같은 문제는 우리에게 큰 우려를 안겨주고 있습니다. 이러한 맥락에서 인공지능이 전력 사용량을 어떻게 증가시키고 있는지에 대한 궁금증이 생겼습니다. 이 궁금증을 해결하기 위해, 선형회귀 분석을 통해 미래의 전력 사용량 추이를 예측하고자 이 주제를 선택하게 되었습니다. 이를 통해 인공지능의 발전이 에너지 소비에 미치는 영향을 더 잘 이해하고, 지속 가능한 발전을 위한 대안을 모색하고자 합니다.

<내용>

- 인공지능의 발전으로 인한 전력사용량 증가를 수학적으로 예측하기 위해서는 2000년부터 2023년까지의 자료를 바탕으로 이후의 전력 사용량 추이를 분석해야 합니다. 미래를 추측하기 위해서 현재의 데이터를 바탕으로 미래에 대한 추세를 분석할 수 있는 선형회귀 분석을 사용해 보았습니다. 선형회귀는 주어진 데이터에 가장 적합한 추세선을 그리는 방법으로, 데이터의 증가 추이를 예측할 수 있습니다. 여기서 가장 적절한 추세선을 긋기 위해서는 데이터의 오차를 최소화하는 선을 찾아야 합니다. 이 과정을 최적화라고 합니다.

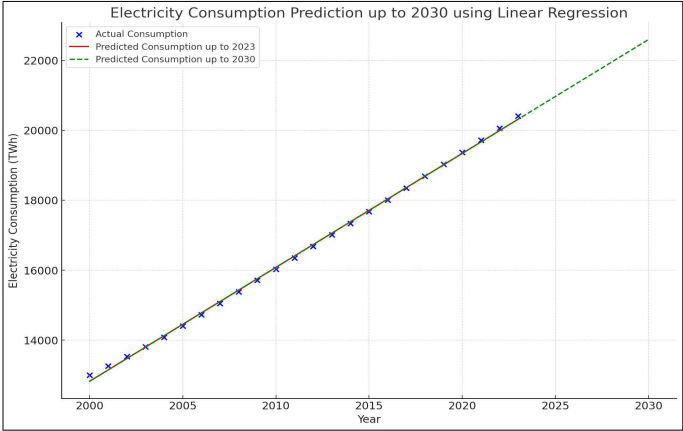
<책내용 및 보고서 작성 내용>

< 최적화는 결국 함수의 극댓값 또는 극솟값을 구하는 문제입니다. 최적화의 대상이 되는 함수를 목적함수라 합니다. 이 목적함수에서 낮춰야 하는 비용이나 시간 등을 구해야 한다면 극솟값을 구하는 문제이고, 높여야 하는 이익이나 성능을 구해야 한다면 극댓값 문제가 됩니다. 따라서 극대 극소 문제는, 모두 동일한 최적화 문제라고 할 수 있습니다. 만약 우리가 최적화 해야하는 목적함수가 우리가 아는 함수로 주어진다면, 그 함수를 미분해서 미분값이 0이 되는 지점에서 도함수의 부호가 바뀌는지를 확인하면 됩니다. 하지만, 현실에서는 목적함수가 우리가 알고 있는 함수 꼴로 친절하게 주어지지 않고, 주어진 x값에 대해서 함숫값 데이터를 구할 수는 있을지언정 수학적인 함수식으로 주어지는 경우는 거의 없습니다. 이런 경우, 극값을 구하는데는 2가지 방법이 존재합니다. 첫 번째 방법은 아무 x값이나 여럿 넣어보면서 가장 큰 결괏값 혹은 가장 작은 결괏값을 찾는 것입니다. 하지만 x값을 전구간에 걸쳐서 조금씩 증가시키는 순차적 증가 방법이나, 구간을 황금 분할하는 첫 번째 방법은 매우 시간이 오래 걸리기 때문에 두 번째 방법인 뉴턴의 미분을 주로 이용합니다. 뉴턴의 미분은 원래 근을 구하는 수치해석 방법입니다. 근 구하기

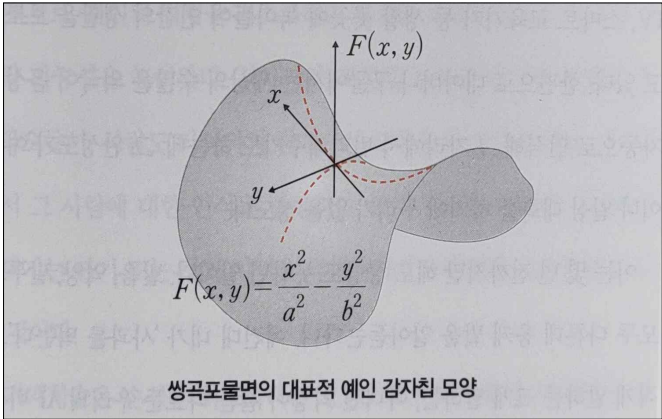
는 fx 가 0이 되는 점을 찾는 일이고, 극점 구하기는 도함숫값 $f'x$ 가 0이 되는 점을 찾는 일입니다. 따라서 뉴턴의 미분을 활용하는 방법은 도함수의 근을 구하는 과정에 함수의 근을 구하는 뉴턴의 방법을 응용한 것이며, 예를 들어 100p의 fx 의 극소점을 찾는 문제에서 도함수 위의 아무 점 x_1 에서 이 점에서의 곡선의 접선을 긋고, 이 접선의 x 절편 x_2 에서 다시 접선을 긋는 작업을 반복하면, 빠르게 $f'x$ 가 x 축과 만나는 점을 찾을 수 있습니다. 최적화 기법은 전통적인 공학 분야에서 최적 설계를 위한 방법으로 널리 활용 되어왔습니다. 그리고 이제는 빅데이터를 대상으로 인식, 비교, 분류, 탐색, 추론 등 복잡한 AI를 학습시키는 핵심 기술로 널리 쓰이고 있습니다. 인공지능의 학습은 기계를 학습시킨다고 해서 기계학습 또는 머신러닝이라고 하는데, 이는 주어진 데이터를 가장 잘 대표하는 최적화된 학습 모델 설계를 의미합니다. 예를 들어보면, 손 글씨를 보고 문자로 인식하고, 얼굴 사진을 보고 누군지 판별하는 것들이 있습니다. 학습이란 사람이나 동물이 직간접적인 경험이나 훈련을 통해서 스스로 지각하고 인지하면서, 자신의 행동이나 인식을 변화시키는 과정입니다. 이 과정에서 자신의 위험을 최소화하고 보상을 최대화함으로써 생명 보존의 가능성을 높여왔습니다. 이와 마찬가지로 기계학습은 손실을 최소화하는 방향으로 이루어집니다. 따라서 손실함수라고 불리는 목적함수를 미분하여, 함수의 극대, 극소값을 구함으로 최적화를 할 수 있습니다. >

저는 최적화 기법 중에서 가장 보편적으로 쓰이는 방법 중 하나인 경사하강법을 사용하였습니다. 경사하강법은 오차의 이계도함수가 0이 되는 지점을 찾는 방법으로, 이를 통해 가장 안정된 상태의 추세를 얻을 수 있습니다.

밑에 있는 그래프는 2000~2023년까지의 전세계의 연간 전력 소비량과 이를 통해 추정한 2024 ~ 2030년의 전력 소비량을 보여줍니다.



위 그래프에서 선형회귀로 추정한 2024~ 2030년의 전력 소비량 추이에서는 인공지능으로 인해 전력 사용량이 증가한 것을 체감하기 힘듭니다. 그 이유는 인공지능으로 인해 증가한 전력소비량은 여러 변수에 기인하기 때문에, 전력 사용량 추이를 정확하게 예측하기 위해서는 다변수 분석이 필요합니다. 이 다변수 분석을 빠르고 정확하게 처리하기 위해서는 미적분 교과과정에서 배우는 일반적인 미분법으로는 한계가 존재하게 됩니다. 이것이 편미분을 도입해야 하는 이유입니다. 예를 들어서 밑의 그림과 같은 쌍곡포물선 모양으로 데이터의 산점도가 분포한다면 x 축에서는 원점이 최솟값, y 축에서는 원점이 최댓값을 갖게 됩니다. 이와 같은 다변수 데이터를 최적화하기 위해서 편미분이 필요합니다.



그럼 편미분은 어떻게 계산하는지 설명해보겠습니다. $f(x) = x^2y^2z^2$ 이라는 변수가 3개 있는 함수가 있다고 생각해 보겠습니다. 밑에 있는 예시를 보면 알수 있듯, 편미분은 선택한 한 변수 외의 다른 변수들을 상수로 취급한 후 미분을 진행하는 미분법입니다. 밑에 예시에서 볼 수 있듯이, 편미분이 계산이 훨씬 간단하며, 편리한 것을 볼 수 있습니다.

$f(x) : x^2 y^2 z^2$
① 음함수 미분
$f'(x) = 2x \cdot y^2 \cdot z^2 + x^2 \cdot 2y \cdot \frac{dy}{dx} + x^2 \cdot y^2 \cdot 2z \cdot \frac{dz}{dx} \dots ?$
② 매개변수 미분
$f'(t) = 2x \cdot \frac{dx}{dt} + y^2 \cdot 2x + x^2 \cdot 2y \cdot \frac{dy}{dt} \cdot x^2 + x^2 \cdot y^2 \cdot 2z \cdot \frac{dz}{dt} \dots ?$
③ 편미분
$f'(x) : 2x \cdot y^2 \cdot z^2$ 계산 편리 ↑↑

여러 변수가 포함되어 있는 손실함수, 데이터의 미분을 통해 얻은 각 변수의 미분값을 벡터로 표현한 뒤, 이를 전부 합성한 것을 그래디언트 벡터라고 합니다. 편미분을 통해 그래디언트 벡터의 방향을 구할 수 있으며, 이를 통해 최적화를 진행합니다. 이 그래디언트 벡터를 합성하는 것이 편미분을 통해 2차원 이상의 고차원 데이터의 경사도를 구하는 방법이며, 그래디언트 벡터를 통해 복잡한 데이터 구조에서도 정확한 최적화를 할 수 있습니다. 하지만, 전력 사용량 추이를 정확하게 예측하기 위해서는 많은 데이터를 분석해야 하는데, 모든 데이터를 경사하강법을 사용하여 하나 하나 미분하는 것은 계산량이 너무 많아지기에 현실적이지 않습니다. 이런 문제를 해결하기 위해서 다 변수 확률적 경사하강법을 사용합니다. 확률적 경사하강법은 각 반복 단계에서 무작위로 선택된 일부 데이터 샘플을 사용하여 경사도를 계산합니다. 이를 통해 연산 비용을 줄이면서도 최적화 과정을 수행할 수 있습니다. 이 방법은 특히 대규모 데이터 세트에서 유용합니다.

<탐구 결과>

인공지능 전력 소비량, 인공지능 발전 속도, 전세계 전력 사용량, 데이터 센터 수 총 4개의 변수를 추가한 후 다변수 최적화를 진행하였습니다. 인공지능의 발전으로 인한 전력 소비량이 증가하고 있음으로 이전보다 명확하게 볼 수 있었습니다. 이는 2021년 기준 대한민국이 2년동안 쓰는 전력량과 비슷한 양입니다. 학습에 사용한 연구 데이터는 골드만 삭스에서 2024년 5월 14일에 발표한 레포트에서 가져왔습니다. 편미분과 확률적 경사하강법을 사용하여 구한 다중 변수 최적화 결과 인공지능의 발전으로 인해 2030년 인공지능이 소비할 것으로 추정되는 전력량은 1160TWh로 달할 것입니다. 이는 2021년 기준 한국이 2년 동안 사용한 전력량과 같은 엄청난 양입니다. 그리고 2030년까지 전 세계 전력 소비 증가율 중 데이터 센터가 차지할 비율은 약 0.9%로 매년 증가할 것으로 예측되는 증가율인 2.4% 중 약 40%를 차지합니다.

<결론>

AI의 발전은 필연적으로 전력 사용량의 증가를 초래할 것입니다. 따라서 우리는 인공지능 기술의 발전뿐만 아니라, 이를 지원할 수 있는 친환경적 전력 생산 방법에 대해서도 많은 논의와 발전이 필요합니다. 그리고, 지속 가능한 발전을 위해 전력 소비 효율성을 높이고, 재생 가능 에너지원의 사용을 확대해야 하며, 특히, 인공지능의 발전으로 인해 데이터 센터의 전력 소비가 급격히 증가할 것으로 예상되므로, 데이터 센터의 에너지 효율성을 높이는 기술 개발과 친환경 에너지원의 활용 및 개발에도 집중해야 한다고 생각합니다.

<출처> AI is poised to drive 160% increase in data center power demand – Goldman Sachs, 미적분의 쓸모 – 한화택