

# 깊은 신경망 모델을 이용한 저밀도 행렬 복원

송원근\*, 이정우

\*서울대학교, 서울대학교

\*swg01110@cml.snu.ac.kr, \*jungle@snu.ac.kr

## Sparse Matrix Recovery using Deep Neural Network

Wongun Song\*, Jungwoo Lee

\*Seoul National Univ., Seoul National Univ.

### 요약

저밀도 행렬 복원은 많은 분야에서 활용이 되고 있고 지금까지 활발하게 연구되어 오고 있는 분야이다. 그러나 지금까지는 원행렬이 저 차원의 행렬이라는 가정을 통해서 복원이 이루어져 왔지만 이는 완벽하지 않다. 그래서 우리는 행렬의 원소들이 상관관계를 가지는 확률변수의 측정값일 경우에 대해서 연구를 진행 하였고, 상호연관된 정규분포로부터 샘플링된 데이터를 가지고 실험을 하여서 기존에 일반적으로 쓰이던 방법보다 좋은 성능을 낸다는 사실을 실험적으로 확인하였다.

### I. 서론

저밀도 행렬 복원은 다양한 적용분야를 가진 연구 분야이다. 예를 들어 사용자들이 각각의 아이템에게 내린 평가 정보를 생각해 보면 우리가 모든 사용자가 모든 아이템에 내린 평가 정보에 대해서 알 수 없으므로 이를 행렬로 나타낼 경우 이도 저밀도 행렬이라고 생각 할 수 있다. 이런 경우 저밀도 행렬 복원 알고리즘을 적용할 경우 사용자가 평가를 내리지 않은 아이템에 대해서 어떤 평가를 내릴지에 대해서도 예측을 할 수 있고, 이를 이용하여 사용자가 아직 평가를 내리지 않았지만 좋은 평가를 내릴것으로 예측 되는 아이템을 추천하는 식으로 추천 시스템을 구성 할 수 있다. 그 이외에도 시스템 식별 및 사물 인터넷 장비의 위치 추측에도 활용이 될 수 있는 등 다양한 응용분야를 가지고 있다.

이런 저밀도 행렬 복원에 대한 알고리즘은 주로 저차원 행렬 인수분해 방법론들을 이용해서 수행이된다. 저차원 행렬 인수분해라는 것은 복원 후의 원본 행렬의 차원이 높지 않다는 가정을 하고 만들어지는 알고리즘이다. 이를 추천시스템의 경우를 예로 들어 다시 얘기하면 유저들이 각 아이템에 내린 평가를 벡터로 나타내었을 때 아이템의 수보다 훨씬 적은 저차원의 선형 공간위에 존재할 것이라는 가정을 한다는 것을 의미한다. 이러한 가정위에서  $M \times N$  행렬이라고 했을 때  $M, N$ 보다 작은  $K$ 에 대해서 그 행렬이  $M \times K, K \times N$  행렬의 곱으로 나타내어 진다고 가정하고 주어진 행렬과의 거리를 최소화 하는 방향으로 원행렬을 찾는 방식을 취하고 있다.

그러나 각각의 벡터들이 저차원 선형공간안에 위치하고 있다는 가정은 실제 경우에는 대부분 사실이 아니다. 실제

경우에서는 각각의 벡터들이 저차원 공간에 존재한다고 해도 이 공간이 비선형 공간일 수도 있고, 행렬의 각각의 원소들이 결정적인 값을 가지는 것이 아니라 확률변수의 측정값일 경우도 많다고 볼 수 있다. 따라서 이런 경우에는 원행렬을 저차원 행렬로 선형적인 인수분해를 하는 것만으로는 부족할 수가 있고 더 복잡한 관계를 예측하기 위해서는 비선형적인 방식이 이용될 필요성이 있다고 생각된다.

이에 따라서 우리는 비선형적인 관계를 추가한 새로운 방법을 제시하였고, 이를 인공적으로 생성된 행렬 데이터를 통해서 기존 가정이 성립되지 않는 환경에서는 기존 방법론보다 더 우수한 성능을 보일 수 있다는 사실을 실험적인 방법을 통해서 증명하였다.

### II. 본론

우리의 방법론은 두 가지로 이루어져 있다. 하나는 각각 행과 열에 해당하는 벡터와 각 행과 열에 해당하는 벡터를 입력으로 받아서 행렬의 원소를 예측할 수 있는 깊은 신경망 모델로 이루어져 있다. 그래서 행렬의  $i, j$  번째의 원소의 값을 구하고 싶을 경우에는  $i$ 번째 행에 해당하는 벡터와  $j$ 번째 열에 해당하는 벡터를 이어붙여서 깊은 신경망 모델을 통과시켜서 값을 추측한다. 그리고 이 모델을 훈련을 시킬 때는 예측값과 값이 주어진 원소와들의 값과의 제곱 오차의 평균이 최소가 되도록 학습을 진행시킨다.

우리가 제안한 방법이 기존방법 대비해서 좋은 점은 첫 번째로는 행과 열에 다른 차원을 부여하는 것이 가능하다는 것이다. 저차원 행렬 인수분해 방법의 경우에는 행에 해당하는 벡터와 열에 해당하는 벡터의 내적 연산이 필요하므로 서로 같은 차원의 공간에 있어야 한다. 하지만 행의 개수와 열의 개수는 보통 비슷하지 않고 한쪽이 더 큰

경우가 많다. 이런 경우 같은 차원으로의 대응은 올바른 표현을 나타내는데에 방해가 될 수 있다. 그리고 두 번째 장점은 단순한 내적이 아닌 깊은 신경망을 이용을 하기 때문에 더 복잡하고 비선형적인 관계에 대해서도 우리의 모델이 표현을 할 수 있다는 점이다.

### III. 실험

실험에 사용한 행렬의 각각의 원소들은 무작위 상관관계를 가진 정규분포로부터 샘플링 된 것으로 행렬은 50개의 행과 40개의 열로 이루어 지고 상관관계 행렬은 정규분포로부터 생성하여 실험에 사용하였다. 그리고 샘플링 된 이후에는 학습때에는 30퍼센트만의 원소를 제외하고 나머지는 모르는 것으로 가정을 하고 학습을 진행하였고 이 때 어떤 행이나 열도 적어도 하나의 원소는 주어지게 하고 실험에 사용하였다.

그리고 베이스라인 방법으로는 저차원 행렬 인수분해 방법을 사용하였고 이때 차원수는 10을 학습률은  $1e-3$ 을 사용하였다. 그리고 우리의 방법에서는 차원수는 행과 열 똑같이 10을 사용하고 깊은 신경망의 경우에는 두 개의 숨겨진 층과 각각의 층은 20 너비를 갖도록 하고 학습률을  $2e-2$ 로 학습을 진행하였다.

실험을 해 본 결과 테이블 1과 같이 우리의 방법론이 기존 방법론에 비해서 오차의 절대값을 평균이나 오차의 분산이 더 낮게 나온다는 것을 확인 하였고, 오차의 분포를 그려본 결과 그림 1, 그림 2를 통해 알 수 있듯이 기존의 방법이 상대적으로 더 무거운 꼬리를 가지고 있는 것을 확인 할 수 있었다.

	평균	표준편차
기존방법	-2.5198274	3.1303174
제안된 방법	-2.4906733	1.5194975

테이블 1. 오차의 평균과 표준편차

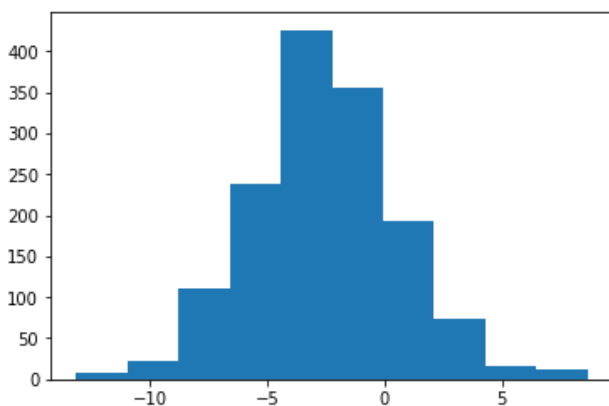


그림1. 기존 방법의 오차의 분포

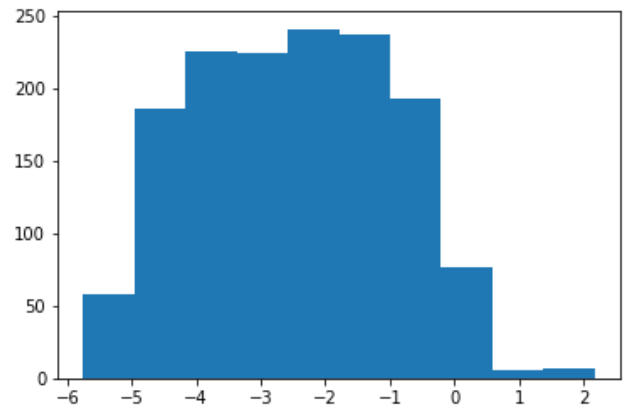


그림2. 제안된 방법의 오차의 분포

### IV. 결론

저차원 가정이 성립되지 않는 행렬복원 문제에서는 기존의 저차원 행렬 인수 분해 방법보다 깊은 신경망을 이용하였을 때 더 좋은 성능을 보일 수 있다는 것을 알 수 있었다

### ACKNOWLEDGMENT

This work is in part supported by Bio-Mimetic Robot Research Center Funded by Defense Acquisition Program Administration, and by Agency for Defense Development (UD190018ID), MSIT-IITP grant (No.2019-0-01367, BabyMind), Grant(UD190031RD) from Defense Acquisition Program Administration(DAPA) and Agency for Defense Development(ADD), INMAC, and BK21-plus.

### 참 고 문 헌

- [1] 류종하, 김영훈, 이혁, 이정우, "Matrix Completion 을 이용한 영화 추천 시스템", *한국통신학회/종합 학술 발표회 논문집*, pp. 286 - 287, Daejeon, Korea, Nov, 2014.
- [2] Candès, E. J., and Recht, B. "Exact Matrix Completion via Convex Optimization". *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, April. 2009.
- [3] Nguyen, L.T., Kim, J., and Shim, B, "Low-Rank Matrix Completion: A Contemporary Survey", *IEEE Access*, vol. 7, no. 1, pp. 94215–94237, July, 2019.
- [4] Candès, E. J., and Plan, Y. (2010). "Matrix Completion with Noise". *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, April, 2010.
- [5] Keshavan, R. H., Montanari, and Oh, S. "Matrix Completion from a Few Entries". *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2980–2998, May, 2010.