# A Feature Selection Study for Medical Image deep learning

Sanghyuck Lee

2021. 05 18.

# Outline

- Deep Feature Selection study
- Thyroid Associated Ophthalmopathy Research
- Industry-University Collaboration Project

# Deep Feature Selection study

$$[X\tilde{X}]^T[X\tilde{X}] = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix} \leftrightarrow \quad \tilde{X} = X(\mathrm{I} - \Sigma^{-1}\mathrm{S}) + \tilde{U}C$$

where $\tilde{U}$ is orthonormal to the column space of $X$, $C \in R^{(n-p)\times p}$ satisfies $\mathrm{C}^{\mathrm{T}}C = 2S - S\,\Sigma^{-1}S \geq 0$, $S$ is $\mathrm{diag}\{s\}$ and $s$=$(s_1, \ldots\, s_p)$ has nonnegative entries.

- Why use a correlation structure

# Deep Feature Selection study

$$[X\tilde{X}]^T[X\tilde{X}] = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix} \leftrightarrow \tilde{X} = X(\mathrm{I} - \Sigma^{-1}\mathrm{S}) + \widetilde{U}C$$

where $\widetilde{U}$ is orthonormal to the column space of $X$, $C \in R^{(n-p)\times p}$ satisfies $\mathrm{C}^{\mathrm{T}}C = 2S - S\,\Sigma^{-1}S \geq 0$, $S$ is diag$\{s\}$ and $s$=$(s_1, \ldots\, s_p)$ has nonnegative entries.

- $X = \begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix}, \Sigma = |1|, \Sigma^{-1} = |1|$

- $\tilde{X} = X(\mathrm{I} - \Sigma^{-1}\mathrm{S}) + \widetilde{U}C = \begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix}(|1| - |1|\mathrm{S}) + \widetilde{U}C$

# Deep Feature Selection study

$$[X\tilde{X}]^T[X\tilde{X}] = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix} \quad \leftrightarrow \quad \tilde{X} = X(\mathrm{I} - \Sigma^{-1}\mathrm{S}) + \widetilde{U}C$$

where $\widetilde{U}$ is orthonormal to the column space of $X$, $C \in R^{(n-p)\times p}$ satisfies $\mathrm{C}^{\mathrm{T}}C = 2S - S\,\Sigma^{-1}S \geq 0$, $S$ is $\mathrm{diag}\{s\}$ and $s$=$(s_1, \ldots\ s_p)$ has nonnegative entries.

- The closer $S$ is to 1, the more significant the feature statistic $W$ is..
- Ex) Lasso coefficients:

If $S$ is close to 0, the lasso coefficient is almost the same, making meaningful comparison difficult.

# Deep Feature Selection study

$$[X\tilde{X}]^T[X\tilde{X}] = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix} \quad \leftrightarrow \quad \tilde{X} = X(\mathrm{I} - \Sigma^{-1}\mathrm{S}) + \tilde{U}C$$

where $\tilde{U}$ is orthonormal to the column space of $X$, $C \in R^{(n-p)\times p}$ satisfies $\mathrm{C}^{\mathrm{T}}C = 2S - S\,\Sigma^{-1}S \geq 0$, $S$ is $\mathrm{diag}\{s\}$ and $s$=$(s_1, \ldots\, s_p)$ has nonnegative entries.

- $[X\tilde{X}]^T[X\tilde{X}] = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix} = G$

  => covariance matrix of augmented design matrix

  => positive semi-definite matrix

  (referenced by Covariance matrix page of Wikipedia)

- 2 methods for calculating $S$

  Semi-Definite Program (SDP)

  Equi-correlated

# Deep Feature Selection study

- $G = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix}$ is positive semi-definite matrix,

  IFF, $2\Sigma - S \geq 0$ and $2S - S\Sigma^{-1}S \geq 0$

Equi-correlated knockoffs:

$s_j = 2\lambda_{min}(\Sigma) \wedge 1$, where $\tilde{X}_j^T X_j = 1 - s_j, \text{diag}\{s\} \leq 2X^T X$

=> minimize $|(\tilde{X}_j^T \quad X_j)|$

# Deep Feature Selection study

$$[X\tilde{X}]^T[X\tilde{X}] = \begin{vmatrix} \Sigma & \Sigma - S \\ \Sigma - S & \Sigma \end{vmatrix} \;\; \leftrightarrow \;\; \tilde{X} = X(\mathrm{I} - \Sigma^{-1}\mathrm{S}) + \tilde{U}C$$

where $\tilde{U}$ is orthonormal to the column space of $X$, $C \in R^{(n-p)\times p}$ satisfies $\mathrm{C}^{\mathrm{T}}C = 2S - S\,\Sigma^{-1}S \geq 0$, $S$ is diag$\{s\}$ and $s$=$(s_1, \ldots\, s_p)$ has nonnegative entries.

- $U^T X = 0$ and $U^T U = I$

  - Deterministic U

    1: Perform QR decomposition $\left[X\ 0_{n\times p}\right] = [Q_x\ Q_0]\,R$

    2: $U \leftarrow Q_0$

# Thyroid Associated Ophthalmopathy Research

- 자동화된 세그멘테이션 모델의 대략적인 성능

- Train: 기존의 세그멘테이션 완료된 AX 3 데이터셋

<table>
<tr><th></th><th>눈</th><th>내직근</th><th>외직근</th><th>시신경</th></tr>
<tr><td>AX 1</td><td rowspan="4">약 90%</td><td>약 50%</td><td>약 50%</td><td>약 50%</td></tr>
<tr><td>AX 2</td><td>약 90%</td><td>약 90%</td><td>약 90%</td></tr>
<tr><td>AX 4</td><td>약 90%</td><td>약 90%</td><td>약 90%</td></tr>
<tr><td>AX 5</td><td>약 50%</td><td>약 50%</td><td>약 50%</td></tr>
</table>

=> 해당 수치에 ×1273 해서 세그멘테이션 완료되었다고 보아도 될듯

# Thyroid Associated Ophthalmopathy Research

- 자동화된 세그멘테이션 모델의 대략적인 성능
- Train: CO 2 – 이번에 수동세그멘테이션 완료한 CO 2 250장

  CO 2 – 이번에 수동세그멘테이션 완료한 CO 3 50장

| | 4 근육 및 시신경 |
|---|---|
| CO 2 | 모두 90% 이상 |
| CO 3 | 전체적으로 실패 |

=> 완료 개수:

CO 2: 대략 250+ 1,023*0.9=1200/1273장 완료, 검토 과정 남음

# Deep Feature Selection study

- 향후 진행해야할 업무:

  - 자동으로 완료한 슬라이스 검토 작업 및 재처리

  - CO3 세그멘테이션

  - AX 1~5종 슬라이스 지방 세그멘테이션

    - 안와 아웃라인 세그멘테이션 이후 내부구조물 제외 작업

  - 완료된 데이터셋 바탕으로 Activity 파일럿 연구

# Industry-University Collaboration Project

- 제출 실적
  - 국제학술대회 논문 1편 이상 게재
  - 국내 특허출원 1건 이상
- 현재 진행 사항
  - 논문 및 특허 주제 선정: 음악 장르 분류
  - 현재까지 진행된 내용:

    제안 방식의 기반이 될 기존 방식 탐색 및 최초 실험

    (조사된 방식: (1) CNN 기반 (2) CRNN 기반)

# References


Barber, Rina Foygel, and Emmanuel J. Candès. "Controlling the false discovery rate via knockoffs." *Annals of Statistics* 43.5 (2015): 2055-2085.

Benjamini, Yoav, and Yosef Hochberg. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." *Journal of the Royal statistical society: series B (Methodological)* 57.1 (1995): 289-300.

Chen, Jiajie, Anthony Hou, and Thomas Y. Hou. "Some analysis of the knockoff filter and its variants." *arXiv preprint arXiv:1706.03400* (2017).

Sarkar, Sanat K., and Cheng Yong Tang. "Adjusting the Benjamini-Hochberg method for controlling the false discovery rate in knockoff assisted variable selection." *arXiv preprint arXiv:2102.09080* (2021).