



20-1 ESC SPRING SESSION

머신러닝에 대한 확률적인 이해

봄학기 세션 소개 및 학술투 모집 공고

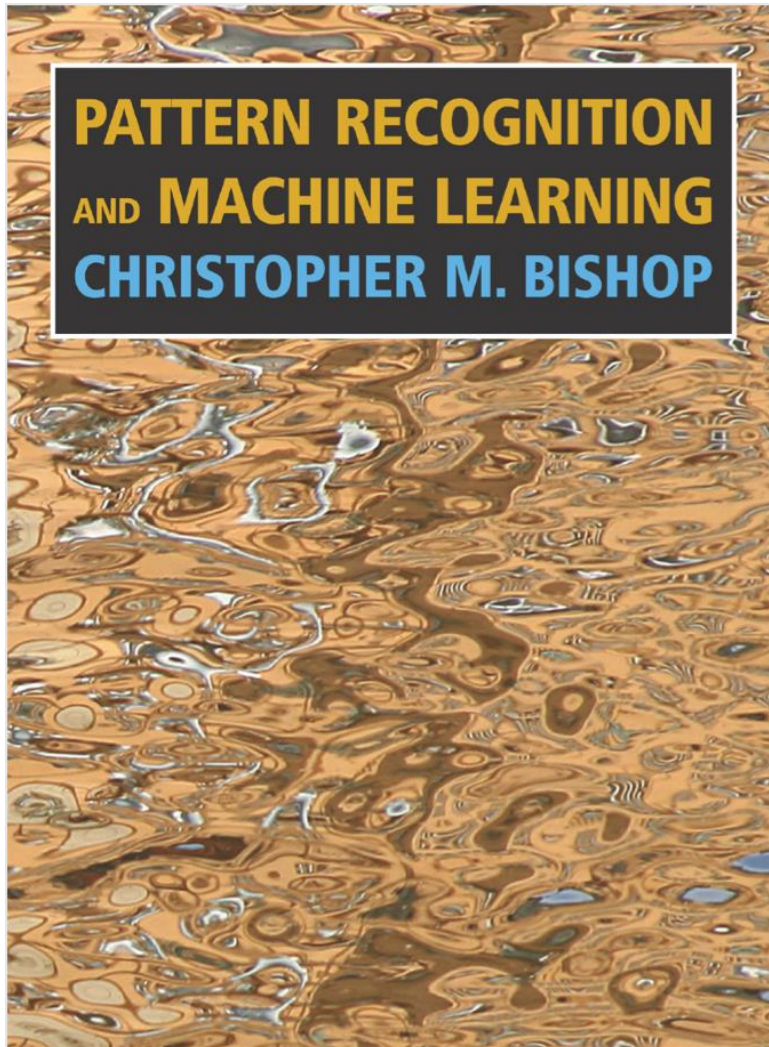
29 MAR 2020

학술투장 강경훈

I. 세션 개요

세션 개요

이번 세션에서 뭐해요?



- *Pattern Recognition and Machine Learning* (2006, Christopher M. Bishop)
- 머신 러닝 교재의 “바이블”
 - ISL이 학부생 수준이라면 PRML은 대학원생 수준
 - 내용이 많이 어려우나 머신러닝의 이론적 이해를 위한 체계를 잡는 데 도움.
 - 학부생 수준에서 어렵거나 불필요한 내용은 거르고 “핵심”만 다룰 예정.
 - 인터넷에서 책의 코드를 구할 수 있음 (<https://github.com/ctgk/PRML>)
그러나 세션에서는 활용하지 않을 것. 실습을 위한 코드는 세션 당일 제공
- 챕터 중 필요한 부분만 골라서 6개 챕터를 6주에 걸쳐 학습,
 - 세션 당일 학술회의 발표 이후 다음 세션 전까지 조별 스터디로 복습 및 과제수행
 - 세션 내용에 해당하는 코딩 과제(필수)와, 심화된 이론 과제(선택)가 나갈 예정
- 마지막 2주는 파이널 프로젝트 기간으로, 실제 Kaggle 등지에서 구한 데이터로 조별로 배운 내용을 활용해 데이터 분석 진행. (자세한 사항은 추후 공지)

세션 개요

세션은 어떻게 진행하나요?



• 매주 세션 구성 (저번 학기의 예시)

Exponential Family: Two-parameter Case

Bayes Rule on a joint prior

ESC

• Prior to Posterior: 4-step ladder process

PRIOR - σ $1/\sigma^2 \sim \text{gamma}(\nu_0/2, \nu_0\sigma_0^2/2)$ $E[\sigma^2] = \sigma_0^2 \frac{\nu_0/2}{\nu_0/2-1}$

- σ_0 는 "내가 생각하는 모집단의 분산", ν_0 는 "그 분산을 계산한 내 마음 속 표본 크기"

PRIOR - μ $\theta|\sigma^2 \sim \text{normal}(\mu_0, \sigma^2/\kappa_0)$

- μ_0 는 "내가 생각하는 모집단의 평균", κ_0 는 "그 평균을 계산한 내 마음 속 표본 크기" (분산 σ_0^2 인 k_0 개의 독립인 확률변수의 합의 평균으로 생각하자)

LIKELIHOOD $Y_1, \dots, Y_n | \theta, \sigma^2 \sim \text{i.i.d. normal}(\theta, \sigma^2)$

POSTERIOR $p(\theta, \sigma^2 | y_1, \dots, y_n) = p(\theta | \sigma^2, y_1, \dots, y_n) p(\sigma^2 | y_1, \dots, y_n)$

MEAN $p(\theta | y_1, \dots, y_n, \sigma^2) \propto p(\theta | \sigma^2) p(y_1, \dots, y_n | \theta, \sigma^2)$

VAR $p(\sigma^2 | y_1, \dots, y_n) \propto p(\sigma^2) p(y_1, \dots, y_n | \sigma^2)$

$= p(\sigma^2) \int p(y_1, \dots, y_n | \theta, \sigma^2) p(\theta | \sigma^2) d\theta$

- Prior와 마찬가지로 Joint posterior가 얻어지므로, 평균과 분산 각각에 대한 사후 분포를 구해 따로 추론을 해야 한다.

FOR INTERNAL USE ONLY - NOT FOR USE AND/OR DISTRIBUTION TO THE GENERAL PUBLIC

14

Graphical representation of the hierarchical model

Arrow는 다음을 가르킨다.

1. parameter \rightarrow random variable

2. hyperparameter \rightarrow parameters

Q1. Hierarchical model에서 hyperprior 왜 선호?

Q2. shared parameter의 장점은?

Q3. 14page 관련이 있다?

회색/흰색의 의미 : 아는 것/모르는 것

동그라미 / 네모 : 확률변수 / 상수

ESC, Yonsei University

16

Hierarchical models for Kdrama rating

1. Overview

우리는 계층모형을 이용하여 한국 드라마 시청률을 살펴봤다.

• Sampling density for group $j = 1, \dots, J$

$Y_{ij} \stackrel{i.i.d.}{\sim} \text{Normal}(\mu_j, \sigma)$

where $i = 1, \dots, n_j$ and n_j is the number of observations in group j

• The stage 1 prior distribution for μ_j

$\mu_j \sim \text{Normal}(\mu, \tau)$

• The stage 2 prior distribution for μ_j

$\mu | \mu_0, \gamma_0 \sim \text{Normal}(\mu_0, \gamma_0)$

$\frac{1}{\tau^2} | \alpha, \beta \sim \text{Gamma}(\alpha, \beta)$

• The prior distribution for σ

$\frac{1}{\sigma^2} | \alpha_\sigma, \beta_\sigma \sim \text{Gamma}(\alpha_\sigma, \beta_\sigma)$

우리는 $\mu_0 = 0.1$ $\gamma_0 = 0.5$ $\alpha_\tau = \beta_\tau = \alpha_\sigma = \beta_\sigma = 1$ 로 뒤서 다음의 결과를 얻었다.

	Lower95	Median	Upper95	Mean	SD	Mode	MCerr	MC.ofSD	SSeff	P
mu	-0.49	0.12	0.66	0.12	0.29	NA	0.00	1.40	5000	0

1. 전 주 내용 Review
혹은 과제 발표
(지원자 자원, 가산점 有)



2. 금주 내용 학술부 발표
(학술부장 및 학술부원)



3. 개별 스터디:
조원이 돌아가면서 복습 티칭
공동으로 과제 수행
(학술부원 참관)



4. FINAL PROJECT

세션 개요

매주 뭐 하면 되나요?



주	일자	전체 세션 주제	발표자	학술부 과제	스터디 조별 과제
1	4/2	OT / Ch01 Introduction	강경훈	Ch03 세션 준비	Ch01 복습 / 과제
2	4/9	Ch03 Linear Models for Regression	강경훈	Ch04 세션 준비	Ch03 복습 / 과제
3	4/16	Ch04 Linear Models for Classification	OOO / XXX	Ch07 세션 준비	Ch04 복습 / 과제
4	4/23	Ch07 Sparse Kernel Machines	OOO / XXX	Ch09 세션 준비	Ch07 복습 / 과제
5	5/7	Ch09 Mixture Models and EM	강경훈 / XXX	Ch12 세션 준비	Ch09 복습 / 과제
6	5/14	Ch12 Continuous Latent Variables	OOO / XXX	Ch14 세션 준비	Ch12 복습 / 과제
7	5/21	Ch14 Combining Models	강경훈 / XXX		Ch14 복습 / 과제
8	5/28	파이널 프로젝트 기간			
9	6/4				

SUN	MON	TUE	WED	THR	FRI	SAT
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
	23기 리쿠르팅					
22	23	24	25	26	27	28
		면접				
29	30	31				

SUN	MON	TUE	WED	THR	FRI	SAT
			1	2 OT 및 1주차	3	4
5	6	7	8	9 2주차	10	11
12	13	14	15	16 3주차	17	18
19	20	21	22	23 4주차	24	25
26	27	28	29	30		

SUN	MON	TUE	WED	THR	FRI	SAT
					1	2
3	4	5	6	7 5주차	8	9
10	11	12	13	14 6주차	15	16
17	18	19	20	21 7주차	22	23
24	25	26	27	28 8주차	29	30
31						

SUN	MON	TUE	WED	THR	FRI	SAT
	1	2	3	4 9주차	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
	기말고사					
28	29	30				

당근

- **우수 조: 과제 제출과 출결 점수가 가장 높은 조에는 전원 기프티콘 증정**
 - 매주 모든 조원이 적시에 과제를 제출하였을 때 해당 조 과제 점수 부여
 - 매주 모든 조원이 세션 당일 (통보하지 않은) 지각/결석이 없을 시 해당 조 출결 점수 부여
 - 학기말 총점이 가장 높은 조 선정
- **우수 학회원: 세션 참여도 (자원 발표, 선택 과제 수행 등)를 고려 1등과 2등에게 6만원/ 3만원 상당의 도서 구입비 지원**
 - 매주 세션 시작 전 전주 내용을 리뷰 발표하거나 선택 과제를 발표하는 경우 가산점 부여
 - 차기 학술부장 선출 시 학술부장의 적극적인 추천!
- **파이널 프로젝트 1등 조: 파이널 프로젝트 1등 시 6만원 상당의 회식비 지원!**

채찍

- **과제를 안 한다, 지각을 한다, 결석을 한다, 등등: 학술부장이 걱정합니다. 많이.. 아주 많이..**

II. 학술부 모집 공고

What is Haksulbu?

많이 빠세요..? 아닙니다!



• 학술부는 뭘 하는가?

- 학술부는 모든 학회원의 원활한 학습을 위해 세션을 준비하고, 각 조에 파견되어 스터디 진행을 돕습니다.
- **학술부 세션:** 학술부는 전체 세션 이전에 학술부 세션을 따로 진행합니다. 학술부 세션에서는 차주 내용을 미리 스터디합니다. 차주 발표를 맡은 학술부원들은 전체 세션 당일까지 발표 자료를 만들고, 출제할 코딩 과제와 이론 과제를 준비합니다. 학술부원은 평균적으로 1번은 2인 1조로 세션 내용을 준비하고, 전체 세션 당일 발표합니다.
 - 세션을 준비하다니 너무 빠센거 아닌가요? 2인 1조이고, 저를 비롯한 모든 학술부원이 매주 세션 내용을 다 공부하고 오기 때문에 부담가지지 않아도 됩니다. 학술부 세션에서는 내가 이해한 바가 맞는지 확인하는 느낌으로 티칭을 하고, 전체 세션에서는 한번 발표하는 경험을 해본다 정도로만 생각하시면 됩니다.
- 때문에 학술부원은 한 주에 1) 학술부 세션 2) 전체 세션 3) 개별 스터디 이렇게 세 번 참여하게 됩니다.
 - 세 번이나 모여요? 어차피 학술부 세션에서 공부 다 하고 가기 때문에 개별 스터디에서 따로 공부할 필요가 없습니다. 관전참ㅋ

[Fig.] Job Description

모집 인원	5명 내외 (현재 4명; 강경훈, 이상욱, 서민희(관전), 오태환(관전))
직무	학술부 세션 (온라인 예정): 차주 내용 스터디 (발표 순번자가 티칭) 전체 세션: (발표 순번자) 세션 준비 및 발표, 과제 출제 개별 스터디: 과제 수행 / 내용 이해 등 스터디 진행 지원
요건	사전지식, 두뇌, 수학 기초 그딴거 필요없고 지원자의 의지 (이왕 온 김에 제대로 배워보겠다!)
혜택	머신러닝에 대한 이해 / 이루 말할 수 없는 보람, 성취감, 자신감 / 학기 종료 후 학술부장이 쓰는 회식

III. 선수 과목

대부분의 머신러닝 방법은 데이터 행렬과 벡터로 이뤄진 Loss function의 최적화(미분)로 구해지기에 다음의 이해가 필요합니다.

1. 선형대수: 다음의 내용을 숙지하고 있으면 됩니다.

- (필수*) 내적: 단위벡터 u 와 벡터 x 의 내적은 u 에서의 x 의 좌표임을 이해하고 있다.
- (필수*) 고유값 분해: Sample Covariance Matrix는 1) 대칭행렬이므로 직교대각화가 가능하고, 2) Positive Definite이므로 고유값이 모두 0 보다 크므로 3) 0이 아닌 양수인 고유값과 고유벡터들의 레이어로 나타낼 수 있다.
- (심화) 특이값 분해: 모든 행렬은 1) input 벡터의 좌표를 다른 좌표체계로 바꾸고, 2) 그 좌표체계에서 변환한 후 output을 또 다른 좌표체계로 바꾸고, 3) 그 결과를 다시 표준좌표체계에서 읽는 식으로 대각화가 가능하다.

2. 벡터미분: 다음을 할 수 있으면 됩니다.

- (필수*) 방향도함수에서 H 가 가지는 의미를 이해하고 있다.
- (필수*) OLS Estimator의 유도: X 을 행렬로, y (target), w (계수벡터)를 벡터로 나타내어 오차를 X 와 y , w 로 이뤄진 norm으로 표현할 수 있으며, 이 식을 w 로 미분하여 최소자승추정량을 산출할 수 있다.
- (삼화) 행렬식의 미분: 행렬식의 미분과 trace 간의 연관성을 알고 있다.

이걸 위해서 수업을 다시 듣거나 교재를 처음부터 끝까지 읽을 필요는 없습니다. 간단한 복습용 강의를 보면서 복기하시면 됩니다.

1. 선형대수 (15강) 유튜브 훈러닝 채널 -> Created Playlists -> 직관적인 선형대수

- (필수*) 2강 내적 / 6강 고유값과 고유벡터 / 7강 대각화 / 8강 유사행렬

 HUN LEARNING	직관적인 선형대수 2. 내적 (Inner Product) 벡터의 길이, 각도, 그리고 영역
 HUN LEARNING	직관적인 선형대수 6. 고유값/벡터 (Eigenvalue/Eigenvector) 선형 변환의 “축”, “스케일”
 HUN LEARNING	직관적인 선형대수 7. 대각화 (Diagonalization) 선형 변환의 “에센스”
 HUN LEARNING	직관적인 선형대수 8. 유사 행렬 (Similar Matrices) 결국 그 놈이 그 놈이다

2. 벡터미분 (4강) 유튜브 훈러닝 채널 -> Created Playlists -> 벡터 및 행렬 미분

- (필수*) 1강 방향도함수 / 2강 벡터미분법칙

 HUN LEARNING	벡터/행렬미분 1. 방향도함수 (Directional Derivative) 어느 방향으로 가는지 알아야지
 HUN LEARNING	벡터/행렬미분 2. 벡터미분법칙 최소자승추정량 유도하기 그냥 알아두면 참 유용한 공식들