

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Суворов Антон Дмитриевич  
Должность: Ректор  
Дата подписания: 27.06.2025 14:26:25  
Уникальный программный ключ:  
a39bdb15d680d3b0adb1ced0af5e1efb14747dc0

## СКОЛКОВСКИЙ ИНСТИТУТ НАУКИ И ТЕХНОЛОГИЙ (Сколтех)

Рабочая программа  
дисциплины

Байесовские методы машинного обучения - 2

Преподаватель

Бурнаев Евгений Владимирович, профессор,  
д.ф-м.н., доцент

### Аннотация

#### Описание курса

Курс посвящен Байесовскому подходу к решению различных задач машинного обучения и анализа данных. Байесовскому подход лежит в основе решения важных прикладных задач, таких как обратные задачи, задачи выбора модели, и задачи непрерывного (онлайн) обучения.

Наибольшее внимание в курсе уделяется слиянию методов глубокого обучения с элементами байесовского подхода. Это слияние стало возможным благодаря развитию за последние 20 лет ряда методов приближенного вывода.

Современные статьи по машинному обучению часто используют методы приближенного вывода. Основным результатом обучения на курсе является умение читать и воспроизводить соответствующие статьи, а также применять соответствующие методы приближенного вывода для разработки алгоритмов байесовского машинного обучения. Для достижения этой цели в курсе предусмотрены теоретические и практические задания, а также итоговый проект.

Цель теоретических заданий двоякая. Во-первых, их выполнение позволит студентам развить навыки получения явных формул, используемых в работах в основе методов приближенного вывода. При этом, в современной литературе фактически отсутствуют источники, по которым можно было бы изучить основные подходы к получению такого рода явных результатов. Выполнение соответствующих заданий позволит восполнить этот пробел. Во-вторых, некоторые теоретические задачи, решенные в рамках упрощенных предположений, позволяют лучше интуитивно понять теоретические свойства соответствующих байесовских методов. Цель практических заданий проста: реализовать обсуждаемые методы в программном виде и изучить их свойства при применении к реальным данным.

Указанные задания подводят учащихся к курсовым проектам: воспроизведение и обсуждение соответствующих статей с передовых конференций по машинному обучению, и/или реализация собственных научно-исследовательских проектов с применением байесовских методов машинного обучения.

Академический уровень курса	Магистратура Аспирантура
Количество кредитов	6

### Предварительные требования к курсу / рекомендации

Перед этим рекомендуется пройти курс численной линейной алгебры и методов оптимизации, а также курс по машинному обучению и приложениям. Кроме того, мы предполагаем, что слушатель свободно владеет линейной алгеброй, вероятностным и реальным анализом.

Тип оценки - дифференцированная

Отображение оценок в процентах

<b>A:</b>	86
<b>B:</b>	70
<b>C:</b>	56
<b>D:</b>	46
<b>E:</b>	36
<b>F:</b>	0

## 2. Содержание курса

Тема	Краткое содержание	Лекции (час)	Семинары (час)	Лабораторные занятия (час)	Самостоятельная работа (час)
Неточный вывод: Глубокий Байес	Вариационные модели автоматического кодирования - Оценка стохастических градиентов и уменьшение дисперсии - Вариационный вывод с нормализацией потоков. - Вариационный отсев. Домашние задания	4	4	1	10
Неточный вывод: Схеме Подходы	Метрополис - Гастингс. Ленгевен Динамика. Масштабируемый Ланжевен Динамика. Гамильтонова динамика. Домашнее задание	4	4	2	10
Неточный вывод: Диффузионные генеративные модели	- Методы подбора баллов: без шума и с нарезкой. - Сети условных оценок по шуму - Гауссовы модели диффузии - Вероятностные модели диффузии с шумоподавлением. Домашнее задание	10	10	2	21

### 3. Результаты обучения

Результаты обучения в Сколтехе указаны в соответствии со структурой результатов обучения в Сколтехе

#### 1. ФУНДАМЕНТАЛЬНЫЕ ЗНАНИЯ

1.1. Знание математики и естественных наук

#### 2.1. ПОЗНАНИЕ И СПОСОБЫ РАССУЖДЕНИЯ

2.1.1. Аналитическое мышление и решение проблем

#### 2.2. ОТНОШЕНИЕ И ПРОЦЕСС ОБУЧЕНИЯ

2.2.2. Готовность принимать решения в условиях неопределенности

2.2.3. Ответственность, интенсивность, настойчивость, безотлагательность и воля к достижению поставленных целей

#### 2.3. ЭТИКА, СПРАВЕДЛИВОСТЬ И ДРУГИЕ ОБЯЗАННОСТИ

2.3.1. Этические действия, честность и мужество

2.3.2. Социальная ответственность

#### 3.1. КОММУНИКАЦИЯ В МЕЖДУНАРОДНОЙ СРЕДЕ

3.1.1. Коммуникационная стратегия и структура

#### 3.2. КОМАНДНАЯ РАБОТА И ЛИДЕРСТВО

3.2.1. Формирование эффективных команд

#### 3.3. СОТРУДНИЧЕСТВО И ИЗМЕНЕНИЯ

3.3.1. Установление разнообразных связей и сетевого взаимодействия

#### 4.1. ПОНИМАНИЕ ГЛОБАЛЬНОГО СОЦИАЛЬНОГО, ЭКОЛОГИЧЕСКОГО И ДЕЛОВОГО КОНТЕКСТА

4.1.2. Принятие ответственности за устойчивое развитие, включая социальные, экономические, экологические аспекты и условия труда

#### 4.2. ДАЛЬНОВИДНОСТЬ — ИЗОБРЕТЕНИЕ НОВЫХ ТЕХНОЛОГИЙ ПОСРЕДСТВОМ ИССЛЕДОВАНИЙ

4.2.1. Процесс исследования — гипотеза, доказательства и защита

#### 4.3. ВИДЕНИЕ — РАЗРАБОТКА КОНЦЕПЦИИ И ПРОЕКТИРОВАНИЕ УСТОЙЧИВЫХ СИСТЕМ

4.3.1. Выявление потребностей и пожеланий заинтересованных сторон

#### 4.4. РЕАЛИЗАЦИЯ КОНЦЕПЦИИ — ВНЕДРЕНИЕ И ЭКСПЛУАТАЦИЯ

4.4.1. Проектирование и оптимизация устойчивого и безопасного внедрения и эксплуатации

#### 4.5. РЕАЛИЗАЦИЯ КОНЦЕПЦИИ — ПРЕДПРИНИМАТЕЛЬСТВО И ПРЕДПРИИМЧИВОСТЬ НА ПРЕДПРИЯТИИ

4.5.1. Разработка концепции и создание нового предприятия

### 4. Задания и выставление оценок

Требование к физической посещаемости (% от числа занятий)	80
---	----

К требованиям о личном посещении:

студентам разрешается пропускать занятия только по уважительной причине, имеющей обоснование, например, медицинское заключение

Тип назначения	Краткое содержание задания	% от итоговой оценки за курс
----------------	----------------------------	------------------------------

Командный проект	Участники курса должны выбрать один проект и выполнить предложенное задание.	30
Домашние задания	Курс состоит из 4 домашних заданий. Каждое домашнее задание состоит из теоретических вопросов и задач по программированию.	40
Промежуточный экзамен	Промежуточный экзамен – это один из тестов в середине курса. Тест состоит из вопросов с несколькими вариантами ответов и коротких задач.	30

## 5. Критерии оценки

<b><u>За д а н и е 1 Т и п а</u></b>	Командный проект
--------------------------------------	------------------

### Пример задания 1

Примеры тем:

1. Глубокие ядра и гауссовы процессы
2. Байесовское активное обучение
3. Ускорение диффузионных моделей
4. Глубокие вариационные автоэнкодеры
5. Байесовское определение точки изменения
6. Сравнение подходов к аппроксимации неразрешимых байесовских моделей
7. МСМС для байесовского вывода
8. Диффузионно-порождающие модели

Итоговый курсовой проект (группы до 3 человек):

- Возможно объединение с проектами из параллельных курсов
- Темы проектов по умолчанию будут объявлены на 3-й неделе
- Этапы: Проектное предложение (неделя 4),
- Проверка статуса этапа (неделя 6),
- Презентация и подача окончательного отчета (неделя 8)

Конечные типы проектов

- Прикладные: выберите интересное приложение и выясните, как применить алгоритмы машинного обучения для его решения;
- Алгоритмические: предложите новый алгоритм обучения или вариант какого-либо существующего для решения общей проблемы или их группы.

Итоговый отчет представлен в формате PDF:

- Введение: мотивация и постановка проблемы
- Сопутствующая работа и краткий обзор литературы
- Описание набора данных

- Методы и алгоритмы ML, предлагаемые модификации алгоритмов и т.д.
  - Эксперименты/обсуждение: подробные сведения о (гиперпараметрах) и о том, как вы их выбирали, метрики и детали перекрестной проверки, обсуждение неудач и успехов, уравнения, результаты, визуализации, таблицы и т.д.
  - Заключение, ссылки, благодарности и вклады.
  - Не более 5 страниц, включая рисунки, таблицы, приложения (только для алгоритмических проектов) и исключая ссылки/ вклады
  - Исходный код (скрипты, блокноты) в формате ZIP или на Github
- Или загрузите образец задания 1

### Пример задания 1

Критерии оценки для задания 1

Основные критерии оценки:

- Общие критерии оценки отчета
  - значимость, новизна: игрушка/реальная проблема или распространенный/неизученный метод
  - техническое качество: продуманный выбор разумных методов, перекрестная проверка и общая оценка качества используемых инструментов/методов
  - общее качество и структура отчета
  - соответствие темам, затронутым в ходе курса
- Презентация проекта
  - качество и ясность презентации
  - актуальное техническое содержание и краткое изложение знаний, продемонстрированных командой

<b><u>Задание 2 Типа</u></b>	Промежуточный экзамен
------------------------------	-----------------------

### Пример задания 2

В середине курса проводится один тест (промежуточный экзамен)

Тест состоит из вопросов с несколькими вариантами ответов и коротких задач.

Основные критерии оценки:

- правильные ответы на вопросы с несколькими вариантами ответов
- полные и корректные решения задач

<b><u>Задание 3 Типа</u></b>	Домашние задания
------------------------------	------------------

### Пример задания 3

Пример задач для домашнего задания:

- 1) Используйте смесь Бернулли для моделирования рукописных цифр из базы

данных рукописных цифр MNIST. Здесь преобразуем цифру в двоичный вектор, установив для всех элементов, значения которых превышают 0,5-1, а для остальных элементов - 0. Применяем смешанную модель для кластеризации рукописных цифр. Исследуем эффективность кластеризации в зависимости от параметров алгоритма.

2) Расширить вариационную трактовку байесовской линейной регрессии, включив в нее гамма-гиперприоритет  $\text{Gam}(\beta|c_0, d_0)$  (по параметру точности шума  $\beta$ ), помимо гиперприоритета гаммы по  $\alpha$  (точность гауссовского априорного распределения по параметрам модели линейной регрессии), предполагая факторизованное вариационное распределение вида  $q(w)q(\alpha)q(\beta\text{-версия})$ . Выведите вариационные уравнения для трех факторов вариационного распределения, а также получите выражение для нижней границы и для прогнозируемого распределения.

### Критерии оценки для задания 3

Решением домашнего задания является:

- отчет в формате PDF о решении задачи с полными, правильными и понятными пояснениями
- исходный код (скрипты, тетради) в ZIP или на Github

Основные критерии оценки:

- полные и правильные решения задач

## **6. Учебники и интернет-ресурсы**

Необходимые учебники	ISBN-13 (or ISBN-10)
Bishop, C.M. Pattern Recognition and Machine Learning. Springer, 2007	978-0387310732
Barber, D. Bayesian Reasoning and Machine Learning. Cambridge University Press, 2012	978-0521518147
Рекомендуемые учебники	
Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009.	9780387848570
Peter Muller, Fernando Andres Quintana, Alejandro Jara, Tim Hanson. Bayesian Nonparametric Data analysis. Springer, 2015	9783319189673
A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin. Bayesian Data Analysis. Chapman & Hall/CRC, 2004	9781439840955
J. Wakefield. Bayesian and Frequentist Regression Methods. Springer, 2013	9781441909244
J.K. Ghosh, M. Delampady, T. Samanta. An Introduction to Bayesian Analysis. Theory and Methods. Springer, 2006	9780387400846
J.K. Ghosh, R.V. Ramamoorthi. Bayesian Nonparametrics. Springer, 2003	9780387955377
M.-H. Chen, Q.-M. Shao, J.G. Ibrahim. Monte Carlo Methods in Bayesian Computation. Springer, 2000	9781461270744
E.G. Phadia. Prior Processes and Their Applications. Springer, 2013	9783642392795
C.P. Robert. The Bayesian Choice. From Decision-Theoretic Foundations to Computational Implementation. Springer, 2007	9780387715988

Веб-ресурсы (ссылки)	Описание
<a href="http://gaussianprocess.org/gpml/">http://gaussianprocess.org/gpml/</a>	Carl Rasmussen and Christopher Williams. Gaussian Processes for Machine Learning. The MIT Press, 2006.
<a href="http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Bqml.HomePage">http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Bqml.HomePage</a>	online version of Barber's "Bayesian Reasoning and Machine Learning"
<a href="http://wol.ra.phy.cam.ac.uk/mackay/itila/book.html">http://wol.ra.phy.cam.ac.uk/mackay/itila/book.html</a>	D. J. C. MacKay (2003) Information Theory, Inference, and Learning Algorithms.

## 7. Оборудование

Программное обеспечение
Требования к программному обеспечению: Python 3.6 и PyTorch: последняя версия NumPy: последняя версия

Оборудование
Доступ к Интернету осуществляется через компьютерный класс и сеть Wi-Fi института.

## Лаборатории для обучения

Лаборатория искусственного интеллекта и суперкомпьютеров  
Лаборатория вычислительной визуализации  
Лаборатория искусственного интеллекта для проектирования материалов  
Лаборатория вычислительного интеллекта  
Лаборатория интеллектуальной обработки сигналов и изображений  
Лаборатория математических основ искусственного интеллекта  
Лаборатория тензорных сетей и глубокого обучения для Интеллектуального Анализа Данных  
Лаборатория “Многомасштабная нейродинамика для интеллектуальных систем”  
Лаборатория мобильной робототехники  
Лаборатория обработки естественного языка  
Лаборатория статистического машинного обучения

## 8. Дополнительные примечания

Комментарии

Материалы/учебное пособие, специально подобранные для различных разделов курса

### 1. Точный вывод и GLM

- S. AMARI and A. CIHOCKI. Information geometry of divergence functions. url: <http://fluid.ippt.gov.pl/bulletin/%2858-1%29183.pdf>
- Rina Foygel Barber, Mathias Darton, and Kean Ming Tan. “Laplace Approximation in Highdimensional Bayesian Regression”. In: arXiv:1503.08337 [math, stat] (Mar. 2015). arXiv: 1503.08337. url: <http://arxiv.org/abs/1503.08337>
- Bradley Efron. “The Geometry of Exponential Families”. EN. In: The Annals of Statistics 6.2 (Mar. 1978), pp. 362–376. issn: 0090-5364, 2168- 8966. doi: 10.1214/aos/1176344130. url: <https://projecteuclid.org/euclid.aos/1176344130>
- Eric P. Xing. The Exponential Family and Generalized Linear Models. url: [https://www.cs.cmu.edu/~epxing/Class/10708-14/scribe\\_notes/scribe\\_note\\_lecture6.pdf](https://www.cs.cmu.edu/~epxing/Class/10708-14/scribe_notes/scribe_note_lecture6.pdf)
- Andrew Gelman and Jennifer Hill. Data analysis using regression and multilevel/hierarchical models. English. OCLC: 646068240. Cambridge; New York: Cambridge University Press, 2007. isbn: 978-0-511-76955-9 978-0- 511-26811-3 978-0-511-79094-2 978-1-282-65283-5. url: <http://dx.doi.org/10.1017/CBO9780511790942>
- S. Kullback and R. A. Leibler. “On Information and Suciency”. EN. In: The Annals of Mathematical Statistics 22.1 (Mar. 1951), pp. 79–86. issn: 0003-4851, 2168-8990. doi: 10.1214/aoms/1177729694. url: <https://projecteuclid.org/euclid.aoms/1177729694>
- M. Tipping M. Bishop. Bayesian Regression and Classification. url: <https://www.microsoft.com/enus/research/wp-content/uploads/2016/02/bishop-nato-bayes.pdf>

- Frank Nielsen and Vincent Garcia. “Statistical exponential families: A digest with flash cards”. In: arXiv:0911.4863 [cs] (Nov. 2009). arXiv: 0911.4863. url: <http://arxiv.org/abs/0911.4863>
- Erlis Ruli, Nicola Sartori, and Laura Ventura. “Improved Laplace Approximation for Marginal Likelihoods”. In: Electronic Journal of Statistics 10.2 (2016). arXiv: 1502.06440, pp. 3986–4009. issn: 1935-7524. doi: 10.1214/16-EJS1218. url: <http://arxiv.org/abs/1502.06440>
- STA 250: Statistics. Laplace Approximation to the Posterior. url: <http://www2.stat.duke.edu/~st118/sta250/laplace.pdf>
- Michael E. Tipping. Bayesian Inference: An introduction to Principles and Practice in Machine Learning. url: [https://www.ics.uci.edu/~smyth/courses/cs274/readings/bayesian\\_regression\\_overview.pdf](https://www.ics.uci.edu/~smyth/courses/cs274/readings/bayesian_regression_overview.pdf)

## 2. Неточный вывод: Вариационные методы

- Shun-ichi Amari. Natural Gradient Works Efficiently in Learning.
- David M. Blei, Alp Kucukelbir, and Jon D. McAulie. “Variational Inference: A Review for Statisticians”. In: Journal of the American Statistical Association 112.518 (Apr. 2017). arXiv: 1601.00670, pp. 859–877. issn: 0162-1459, 1537-274X. doi: 10.1080/01621459.2017.1285773. url: <http://arxiv.org/abs/1601.00670>
- EM and MM Algorithms. url: [https://www.samsi.info/wp-content/uploads/2016/08/Zhou\\_samsi-optsummerschool-20160810-1.pdf](https://www.samsi.info/wp-content/uploads/2016/08/Zhou_samsi-optsummerschool-20160810-1.pdf)
- Matt Hoffman et al. “Stochastic Variational Inference”. In: arXiv:1206.7051 [cs, stat] (June 2012). arXiv: 1206.7051. url: <http://arxiv.org/abs/1206.7051>
- M. Tipping M. Bishop. Mixtures of Probabilistic PCA. url: <http://www.miketipping.com/papers/metmppca.pdf>
- Razvan Pascanu and Yoshua Bengio. “Revisiting Natural Gradient for Deep Networks”. In: arXiv:1301.3584 [cs] (Jan. 2013). arXiv: 1301.3584. url: <http://arxiv.org/abs/1301.3584>
- Michael E. Tipping. Probabilistic PCA. url: <https://www.microsoft.com/en-us/research/wpcontent/uploads/2016/02/bishop-ppca-jrss.pdf>
- Tong Tong Wu and Kenneth Lange. “The MM Alternative to EM”. In: Statistical Science 25.4 (Nov. 10 2010). arXiv: 1104.2203, pp. 492–505. issn: 0883-4237. doi: 10.1214/08-STS264. url: <http://arxiv.org/abs/1104.2203>
- Anderson Y. Zhang. Theoretical and computational guarantees of mean field variational inference for community detection. url: <http://www.stat.yale.edu/~hz68/variational.pdf>

## 3. Неточный вывод: Глубокое обучение и байесовские методы

- Martin Arjovsky, Soumith Chintala, and Leon Bottou. “Wasserstein GAN”. In: arXiv:1701.07875 [cs, stat] (Jan. 2017). arXiv: 1701.07875. url: <http://arxiv.org/abs/1701.07875>
- Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. “Importance Weighted Autoencoders”. In: arXiv:1509.00519 [cs, stat] (Sept. 2015). arXiv: 1509.00519. url: <http://arxiv.org/abs/1509.00519>

- Chris Cremer, Quaid Morris, and David Duvenaud. “Reinterpreting Importance-Weighted Autoencoders”. In: arXiv:1704.02916 [stat] (Apr. 2017). arXiv: 1704.02916. url: <http://arxiv.org/abs/1704.02916>
- Alex Graves. “Stochastic Backpropagation through Mixture Density distributions”. In: arXiv:1607.05690 [cs] (July 2016). arXiv: 1607.05690. url: <http://arxiv.org/abs/1607.05690>
- Shixiang Gu et al. “MuProp: Unbiased Backpropagation for Stochastic Neural Networks”. en. In: (Nov. 2015). url: <https://arxiv.org/abs/1511.05176>
- Jiri Hron, Alexander G. de G. Matthews, and Zoubin Ghahramani. “Variational Gaussian Dropout is not Bayesian”. In: arXiv:1711.02989 [stat] (Nov. 2017). arXiv: 1711.02989. url: <http://arxiv.org/abs/1711.02989>
- Eric Jang, Shixiang Gu, and Ben Poole. “Categorical Reparameterization with Gumbel-Softmax”. In: arXiv:1611.01144 [cs, stat] (Nov. 2016). arXiv: 1611.01144. url: <http://arxiv.org/abs/1611.01144>
- Diederik P. Kingma, Tim Salimans, and Max Welling. “Variational Dropout and the Local Reparameterization Trick”. In: arXiv:1506.02557 [cs, stat] (June 2015). arXiv: 1506.02557. url: <http://arxiv.org/abs/1506.02557>
- Diederik P. Kingma and Max Welling. “Auto-Encoding Variational Bayes”. In: arXiv:1312.6114 [cs, stat] (Dec. 2013). arXiv: 1312.6114. url: <http://arxiv.org/abs/1312.6114>
- Lars Maaløe et al. “Auxiliary Deep Generative Models”. In: arXiv:1602.05473 [cs, stat] (Feb. 2016). arXiv: 1602.05473. url: <http://arxiv.org/abs/1602.05473>
- Andriy Mnih and Karol Gregor. “Neural Variational Inference and Learning in Belief Networks”. en. In: (Jan. 2014). url: <https://arxiv.org/abs/1402.0030>
- Dmitry Molchanov, Arsenii Ashukha, and Dmitry Vetrov. “Variational Dropout Sparsifies Deep Neural Networks”. In: arXiv:1701.05369 [cs, stat] (Jan. 2017). arXiv: 1701.05369. url: <http://arxiv.org/abs/1701.05369>
- Sebastian Nowozin, Botond Cseke, and Ryota Tomioka. “f-GAN: Training Generative Neural Samplers using Variational Divergence Minimization”. In: arXiv:1606.00709 [cs, stat] (June 2016). arXiv: 1606.00709. url: <http://arxiv.org/abs/1606.00709>
- Mihaela Rosca et al. “Variational Approaches for Auto-Encoding Generative Adversarial Networks”. In: arXiv:1706.04987 [cs, stat] (June 2017). arXiv: 1706.04987. url: <http://arxiv.org/abs/1706.04987>
- Francisco J. R. Ruiz, Michalis K. Titsias, and David M. Blei. “The Generalized Reparameterization Gradient”. In: arXiv:1610.02287 [stat] (Oct. 2016). arXiv: 1610.02287. url: <http://arxiv.org/abs/1610.02287>
- John Schulman et al. “Gradient Estimation Using Stochastic Computation Graphs”. In: arXiv:1506.05254 [cs] (June 2015). arXiv: 1506.05254. url: <http://arxiv.org/abs/1506.05254>
- Jakub M. Tomczak and Max Welling. “VAE with a VampPrior”. In: arXiv:1705.07120 [cs, stat] (May 2017). arXiv: 1705.07120. url: <http://arxiv.org/abs/1705.07120>
- George Tucker et al. “REBAR: Low-variance, unbiased gradient estimates for discrete latent variable models”. In: (Mar. 2017).

#### 4. Неточный вывод: методы MCMC

- Handbook of Markov Chain Monte Carlo. url: <http://www.mcmchandbook.net/>
- Yi-An Ma, Tianqi Chen, and Emily B. Fox. “A Complete Recipe for Stochastic Gradient MCMC”. In: arXiv:1506.04696 [math, stat] (June 2015). arXiv: 1506.04696. url: <http://arxiv.org/abs/1506.04696>
- Monte Carlo theory, methods and examples. url: <http://statweb.stanford.edu/~owen/mc/>
- Andriy Mnih Ruslan Salakhutdinov. Bayesian Probabilistic Matrix Factorization using Markov Chain Monte Carlo. url: <https://www.cs.toronto.edu/~amnih/papers/bpmf.pdf>
- Yee Whye Teh Welling. Bayesian Learning via Stochastic Gradient Langevin Dynamics. url: [https://www.ics.uci.edu/~welling/publications/papers/stoclangevin\\_v6.pdf](https://www.ics.uci.edu/~welling/publications/papers/stoclangevin_v6.pdf)

## 5. Гауссовский процесс и вариационная оптимизация

- Ryan P. Adams A Tutorial on Bayesian Optimization for Machine Learning. url: [https://www.iro.umontreal.ca/~bengioy/cifar/NCAP2014-summer-school/slides/Ryan\\_adams\\_140814\\_bayesopt\\_ncap.pdf](https://www.iro.umontreal.ca/~bengioy/cifar/NCAP2014-summer-school/slides/Ryan_adams_140814_bayesopt_ncap.pdf)
- Mauricio A. Alvarez, Lorenzo Rosasco, and Neil D. Lawrence. “Kernels for Vector-Valued Functions: a Review”. In: arXiv:1106.6251 [cs, math, stat] (June 2011). arXiv: 1106.6251. url: <http://arxiv.org/abs/1106.6251>
- Eduard Gabriel Bazavan. Fourier Kernel Learning. url: <http://www.maths.lth.se/matematiklth/personal/sminchis/papers/eccv12bls.Pdf>
- [David Barber. Optimization by Variational Bounding. url: <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-65.pdf>
- Gaussian Processes for Machine Learning: Book webpage. url: <http://www.gaussianprocess.org/gpml/>
- Gaussian Process and Uncertainty Quantification Summer School 2017. original-date: 2016-12-22T22:57:23Z. July 2018. url: <https://github.com/gpschool/gpss17>
- James Hensman, Nicolas Durrande, and Arno Solin. “Variational Fourier features for Gaussian processes”. In: arXiv:1611.06740 [stat] (Nov. 2016). arXiv: 1611.06740. url: <http://arxiv.org/abs/1611.06740>
- Mehryar Mohri. Optimistic Bandit Convex Optimization. url: <https://cs.nyu.edu/~mohri/pub/bcop.pdf>
- Tim Salimans et al. “Evolution Strategies as a Scalable Alternative to Reinforcement Learning”. In: arXiv:1703.03864 [cs, stat] (Mar. 2017). arXiv: 1703.03864. url: <http://arxiv.org/abs/1703.03864>
- Bobak Shahriari. Taking the Human Out of the Loop: A Review of Bayesian Optimization. url: <https://www.cs.ox.ac.uk/people/nando.defreitas/publications/BayesOptLoop.pdf>
- Joe Staines and David Barber. “Variational Optimization”. In: arXiv:1212.4507 [cs, stat] (Dec. 2012). arXiv: 1212.4507. url: <http://arxiv.org/abs/1212.4507>
- Titsias. Variational Learning of Inducing Variables in Sparse Gaussian Processes. url: <http://proceedings.mlr.press/v5/titsias09a/titsias09a.pdf>
- Tianbao Yang. Nystrom Method vs Random Fourier Features: A Theoretical and Empirical Comparison. url: <https://papers.nips.cc/paper/4588-nystrom-method-vs-random-fourier-features-atheoretical-and-empirical-comparison.pdf>

## 6. Последовательные данные

- Blei. Exact Inference: Elimination and Sum Product (and hidden Markov models). url: <http://www.cs.columbia.edu/~blei/fogm/2015F/notes/inference.pdf>
- Henrik I. Christensen. Graphical Models & HMMs. url: <https://www.cc.gatech.edu/~hic/CS7616/pdf/lecture6.pdf>
- Eric P. Xing. Kalman Filtering. url: [http://www.cs.cmu.edu/~epxing/Class/10708-16/note/10708\\_scribe\\_lecture12.pdf](http://www.cs.cmu.edu/~epxing/Class/10708-16/note/10708_scribe_lecture12.pdf)
- Jonathan S. Yedidia. Understanding Belief Propagation and its Generalizations. url: <http://www.merl.com/publications/docs/TR2001-22.pdf>
- Thomas Minka. “Divergence Measures and Message Passing.” Accessed August 9, 2018. <https://pdfs.semanticscholar.org/40b0/02847cecce2e3cfaeae83e85393e57179a0c.pdf>

## 7. Неточный вывод: диффузионные генеративные модели

- Yang Song. Generative Modeling by Estimating Gradients of the Data Distribution. url: <https://proceedings.neurips.cc/paper/2019/file/3001ef257407d5a371a96dcd947c7d93-Paper.pdf>
- Jonathan Ho. Denoising Diffusion Probabilistic Models. url: <https://deepsense.ai/wp-content/uploads/2023/03/2006.11239.pdf>
- Yang Song. Score-Based Generative Modeling through Stochastic Differential Equations url: <https://openreview.net/forum?id=PXTIG12RRHS>