

Bachelorseminar: Variationsmethoden in der Angewandten Mathematik

Betreuer: Martin Burger, Daniel Tenbrinck, Philipp Wacker

Sprache: Deutsch

Benötigtes Vorwissen:

Grundlagen der linearen Algebra, Analysis und numerische Methoden

Zielgruppe:

→ Studierende ab dem 5. (ev. 3.) Fachsemester der Studiengänge:
*B.Sc. Mathematik / Technomathematik / Wirtschaftsmathematik
bei entsprechenden Vorkenntnissen*

Inhalt:

Die Seminarteilnehmer beschäftigen sich mit aktuellen Themen im Bereich der Variationsmethoden, mit verschiedenen Themen, die zusammen verschiedene Aspekte wie Analysis, Numerik und verschiedenen Anwendungen abdecken. Variationsmethoden basieren auf der Konstruktion geeigneter Energien, entweder aus physikalischen oder rein mathematischen Gründen, deren Minimierung geeignete Lösungen ergibt. Klassische Beispiele sind die Mechanik, in der man Gleichgewichtslagen durch Minimierung potentieller Energien erhält, oder die Maximum-Likelihood Methode in der Statistik. Die Studierenden sollen sich hierbei sowohl den theoretischen Hintergrund als auch die praktische Lösung der vorgestellten Problemstellungen erarbeiten. Beispiele für Themen sind die folgenden Gebiete:

Gradientenflüsse

Gradientenflüsse sind spezielle gewöhnliche Differentialgleichungen, bei denen die rechte Seite der Gradient einer Energie ist, d.h.

$$u'(t) = -\nabla E(u(t))$$

Gradientenflüsse verkleinern im Verlauf die Energie der Lösung und konvergieren unter geeigneten Bedingungen an die Energie E gegen einen Minimierer. Eine interessante Eigenschaft zur Approximation von Lösung ist die variationelle Formulierung des impliziten Euler-Verfahrens

$$\frac{u(t) - u(t - \tau)}{\tau} = -\nabla E(u(t))$$

Äquivalent kann man $u(t)$ als Minimierer von

$$E_\tau(u) = \frac{\|u - u(t - \tau)\|^2}{2\tau} + E(u)$$

charakterisieren und daraus kann man einerseits mit variationellen Methoden Eigenschaften und Konvergenz der Diskretisierung untersuchen und andererseits mit geeigneten Optimierungsverfahren auch numerische Verfahren implementieren.

Literatur:

- Santambrogio, F. (2017). {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bulletin of Mathematical Sciences*, 7(1), 87-154.

Stein Variational Gradient Descent

Im Kontext von Bayesschen inversen Problemen, insbesondere in hohen Dimensionen, ist es oft sinnvoll, ein Wahrscheinlichkeitsmaß durch eine "Punktwolke" von repräsentativen Parametern abzubilden. Eine Möglichkeit dazu ist das Erzeugen von Samples, z.B. mit Monte-Carlo-Methoden. Wir werden eine alternative Methode untersuchen, den Stein Variational Gradient Descent, der mittels einer variationellen Formulierung ein Ensemble von Punkten so bewegt, dass es ein gegebenes Zielmaß gut approximiert.

Literatur:

- Q. Liu, D. Wang: Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm, NIPS 2016
- J. Lu, Y. Lu, J. Nolen, Scaling limit of the Stein variational gradient descent: The mean field regime, SIAM Journal on Mathematical Analysis, 2019
- Q. Liu, Stein Variational Gradient Descent as Gradient Flow, NIPS 2017

Bayessche Inverse Probleme -- Maximum Likelihood, Maximum-a-Posteriori, Sampling

Ein inverses Problem beschäftigt sich mit der Frage, eine Größe zu schätzen, über die wir durch einen fehlerbehafteten und indirekten Messprozess nur begrenzte Information besitzen. Im Bayesschen Zugang zu inversen Problemen versuchen wir, diese unbekannte Größe nicht durch einen festen Wert auszudrücken (wie z.B. beim Least-Squares-Modell), sondern durch eine (Wahrscheinlichkeits-)Verteilung, der a-Posteriori-Verteilung, die durch ihre Lage und Konzentration einen Bereich definiert, in dem der gesuchte Parameter mit großer Wahrscheinlichkeit zu finden ist.

Das größte praktische Problem in diesem Kontext ist meist das Extrahieren von Information aus dem Posterior, also das Berechnen von Mittelwert, Varianz, dem Punktschätzer maximaler Dichte, sowie das Erzeugen von "typischen" Stichproben.

Literatur:

- Allmaras et al., Estimating Parameters in Physical Models through Bayesian Inversion: A Complete Example, SIAM Review
- J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Springer

Randomize-Then-Optimize

RTO ist eine Sampling-Methode, um aus nicht-Gaußschen Wahrscheinlichkeitsmaßen Stichproben (Samples) zu ziehen. Die Idee ist dabei, die Daten auf geschickte Weise so zu perturbieren, dass jeder resultierende Maximum-A-Posteriori-Schätzer dieser perturbierten Daten ein (approximatives) Sample aus der Verteilung darstellt.

Literatur:

- J. Bardsley, A. Solonen, H. Haario, M. Laine, RANDOMIZE-THEN-OPTIMIZE: A METHOD FOR SAMPLING FROM POSTERIOR DISTRIBUTIONS IN NONLINEAR INVERSE PROBLEMS, SIAM J. Sci. Comput.
- J. Bardsley, T. Cui, Y. Marzouk, Z. Wang, SCALABLE OPTIMIZATION-BASED SAMPLING ON FUNCTION SPACE, SIAM J. Sci. Comput.

Poisson-Learning auf Graphen

Graphen erlauben es beliebige Datenstrukturen geeignet zu erfassen und abzubilden, z.B., Bilder, Soziale Netzwerke, oder hochdimensionale Merkmalsvektoren im Maschinellen Lernen. Die Knoten des Graphen repräsentieren die Daten während die Kanten die Beziehung der Daten untereinander modellieren und ihnen ein Gewicht geben.

Eine etablierte Methode des maschinellen Lernens mit Hilfe von Graphen ist das sogenannte Laplacian Learning bei der man ausgehend von einer Handvoll von manuell annotierten Knoten versucht eine Klassifikation auf die unannotierten Daten zu übertragen. Dieses Verfahren basiert auf der Lösung eines Variationsproblems auf dem Graphen und funktioniert gut für eine genügend große Menge von Annotationen. Laplacian Learning funktioniert leider nur unbefriedigend, wenn man nur sehr wenig annotierte Daten vorliegen hat. Hier setzt eine kürzlich vorgeschlagene Verbesserung, das sogenannte Poisson Learning, an mit dessen Hilfe man auch sehr gute Klassifikationen trotz weniger annotierter Daten erhalten kann.

Literatur:

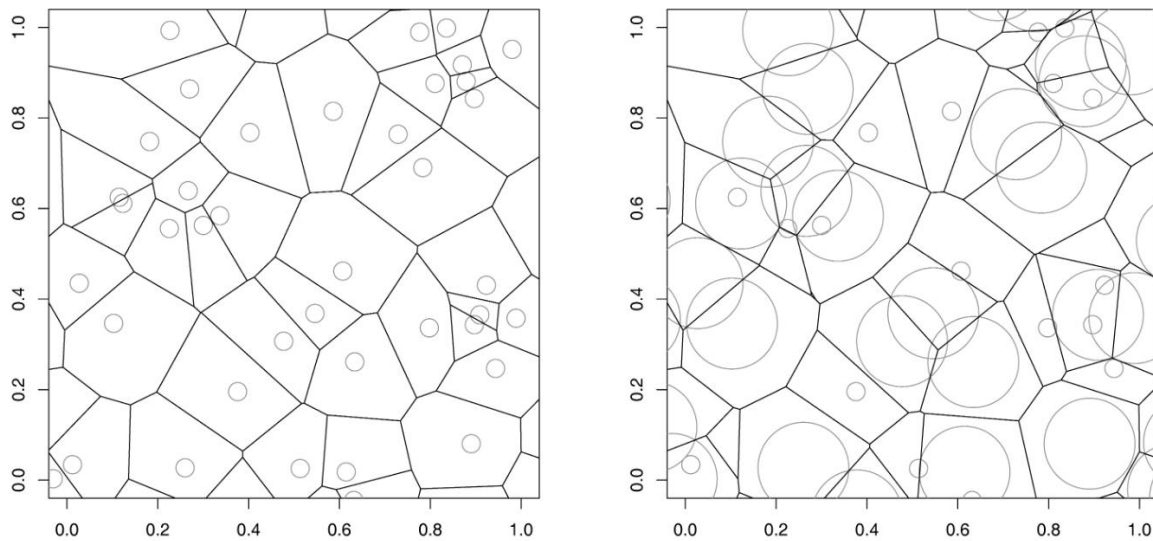
- J. Calder, B. Cook, M. Thorpe, D. Slepcev: Poisson Learning: Graph Based Semi-Supervised Learning At Very Low Label Rates, [arXiv:2006.11184](https://arxiv.org/abs/2006.11184)

Laguerre-Tesselation von Daten für Clustering

Clustering-Methoden zählen zu den unüberwachten Lernmethoden im maschinellen Lernen. Standardalgorithmen, wie z.B. k-means, liefern häufige gute Gruppierungen einer unbekannt Datenmenge in sogenannte Cluster. Die dabei induzierte Partition der Daten ist direkt beschreibbar durch Voronoizellen, die die Geometrie dieser Partition repräsentieren und Entscheidungsgrenzen für Algorithmen der KI darstellen.

Eine Verallgemeinerung von Voronoizellen sind sogenannte Laguerrezellen, die eine Gewichtung eines Clusters erlauben und somit adaptiver zur Datenanalyse genutzt werden können.

In diesem Thema soll es um die Unterschiede zwischen Voronoi- und Laguerrezellen gehen und um ihre Beziehung zu Clustering-Algorithmen.

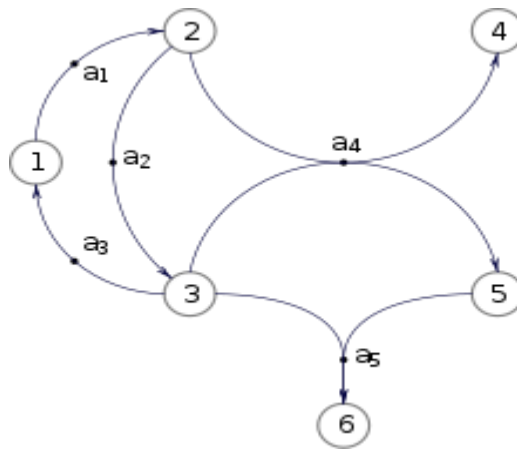


Literatur:

- Q. Duan, D. P. Kroese, T. Brereton, A. Spettl, V. Schmidt: Inverting Laguerre Tessellations, *The Computer Journal* 57(9):1431-1440 (2014)
- C. Lautensack: Random Laguerre Tessellations, Dissertation der TH Karlsruhe (2007)

Variationsmethoden auf Hypergraphen

Hypergraphen sind eine Verallgemeinerung von Graphen, bei denen eine Kante auch zwischen mehr als zwei Knoten führen kann. Solche Strukturen sind von großem Interesse in modernen Datenstrukturen, insbesondere in sozialen Netzwerken, wo Communities und Interaktionen nicht nur aus Paaren bestehen. Nun hat man oft skalare oder vektorielle Daten in den Knoten, etwa den Grad der Zustimmung zu einem bestimmten Thema. Viele Fragestellungen, etwa das Auffinden von Clustern, können nun durch Variationsmethoden auf Hypergraphen gelöst werden. Insbesondere Methoden wie totale Variation sind hier interessant. Dabei wird eine Energie formuliert, die die Ähnlichkeit der skalaren Größe zwischen Knoten, die durch Hyperkanten verbunden sind, quantifiziert. Dabei ist die geeignete Definition solcher Energie ein genauso interessantes Problem wie die numerische Berechnung von Minimierern.



Literatur:

- Hein, M., Setzer, S., Jost, L., & Rangapuram, S. S. (2013). The total variation on hypergraphs - learning on hypergraphs revisited. NIPS 2013