Chapter 0
Organization and Overview

Convex Optimization for Machine Learning & Computer Vision
SS 2017

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Computer Vision Group
Department of Informatics
TU München

Organization

Whether this lecture fits you?

Prerequisites

- Interest in mathematical theory
- Background in analysis and linear algebra
- Numerical implementation (in Matlab or Python)

Whether this lecture fits you?

Prerequisites

- Interest in mathematical theory
- Background in analysis and linear algebra
- Numerical implementation (in Matlab or Python)

Nice plus (but not necessary)

- Experience in machine learning or computer vision
- Knowledge in continuous optimization
- Knowledge in functional analysis

Course overview

Lectures

1. Theory of convex analysis
2. Design and analysis of optimization algorithms
3. Selected topic: Bilevel optimization

Applications are mostly covered by exercise sessions...

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Exercise session

Organizers: Thomas Möllenhoff and Emanuel Laude

- Exercise sheets covering the content of the lecture will be passed out every Wednesday
- Exercises contain theoretical as well as programming problems
- You have one week for the exercise sheets and can turn in your solution
- You may work on the exercises in groups of two
- Reaching at least 60% of the total exercise points results in a 0.3 grade bonus
- If solutions have obviously been copied, both groups will get 0 points

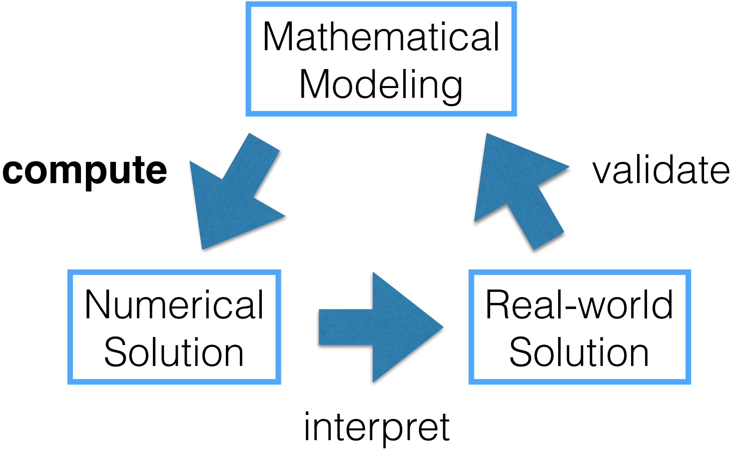
Contact us

Miscellaneous info

- Tao's office: 02.09.061
- Thomas' office: 02.09.060
- Emanuel's office: 02.09.039
- Office hours: Please write an email.
- Lecture: Starts at quarter past. Short break in between.
- Course website:
  `https://vision.in.tum.de/teaching/ss2017/cvx4cv`

First Glimpse of the Course

Driving cycle



An example from computer vision / machine learning

- Image segmentation / multi-labeling:

image

segmentation ($L = 4$)

An example from computer vision / machine learning

- Image segmentation / multi-labeling:

- Variational method for finding label function $u : \Omega \to \Delta^L$

$$\text{minimize} \int_\Omega \Big(\delta\{u(x) \in \Delta^L\} + u(x) \cdot f(x)\Big)\, dx + \sum_{l=1}^{L} \mathrm{TV}_\omega(u^l),$$

where

- Δ^L is the unit $(L-1)$-simplex.
- $f : \Omega \to \mathbb{R}^L$ is the (pre-computed) *unary potential*.
- $\mathrm{TV}_\omega(\cdot)$ is the (weighted) total-variation norm.

Prototypical workflow

- The variational model

$$\text{minimize} \int_{\Omega} \Big(\delta\{u(x) \in \Delta^L\} + u(x) \cdot f(x) \Big)\, dx + \sum_{l=1}^{L} \mathrm{TV}_{\omega}(u^l)$$

  is a special case of **convex optimization**

$$\text{minimize } J(u) + \delta\{u \in C\},$$

  with **convex objective** J and **convex constraint** C.

- This course is about **theory** and **practice** for solving convex optimization (arising from computer vision and machine learning).

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Prototypical workflow

- Put into canonical form:

$$\min_{u} F(Ku) + G(u),$$

where F, G are convex functions, K is a linear operator.

Prototypical workflow

- Put into canonical form:

$$\min_{u} F(Ku) + G(u),$$

where F, G are convex functions, K is a linear operator.

- Reformulate the problem (by introducing *dual variable* p):

$$\max_{p} \min_{u} \langle Ku, p \rangle - F^*(p) + G(u),$$

where F^* is the *convex conjugate* of F.

Tao Wu
Thomas Möllenhoff
Emanuel Laude

Prototypical workflow

- Put into canonical form:

$$\min_u F(Ku) + G(u),$$

where F, G are convex functions, K is a linear operator.

- Reformulate the problem (by introducing *dual variable* p):

$$\max_p \min_u \langle Ku, p \rangle - F^*(p) + G(u),$$

where F^* is the *convex conjugate* of F.

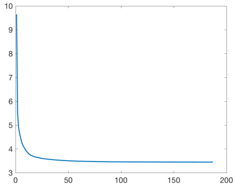
- Apply an iterative scheme (e.g. PDHG):

$$u^{k+1} = \arg\min_u \left\langle u, K^\top p^k \right\rangle + G(u) + \frac{s}{2}\|u - u^k\|^2,$$
$$p^{k+1} = \arg\min_p -\left\langle K(2u^{k+1} - u^k), p \right\rangle + F^*(p) + \frac{t}{2}\|p - p^k\|^2.$$

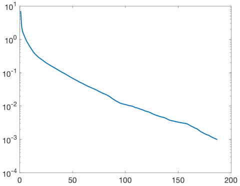
Tao Wu
Thomas Möllenhoff
Emanuel Laude

Questions of our concern

energy value



primal-dual gap



- Does a minimizer always exist?
- How to characterize a minimizer?
- How to derive an iterative solver?
- How to analyze the convergence of the algorithm?
- How to accelerate the convergence?
- Efficient implementation $\rightsquigarrow$ exercise session.

Ready to start?

Tao Wu
Thomas Möllenhoff
Emanuel Laude