Model Predictive Control and Reinforcement Learning
– Lecture 2.2: Basics in Optimization –

Andrea Ghezzi and Moritz Diehl

University of Freiburg

Slides jointly developed with Armin Nurkanović, Florian Messerer, Katrin Baumgärtner.

Fall School on Model Predictive Control and Reinforcement Learning
Freiburg, 6-10 October 2025

Outline of the lecture

Basic definitions

Some classifications of optimization problems

Optimality conditions for unconstrained optimization

Newton's method for unconstrained optimization

What is an optimization problem?

Minimize (or maximize) an objective function $F(w)$ depending on decision variables w subject to equality and/or inequality constrains

What is an optimization problem?

Minimize (or maximize) an objective function $F(w)$ depending on decision variables w subject to equality and/or inequality constrains

An optimization problem

$$\min_{w} \; F(w) \tag{1a}$$

$$\text{s.t.} \quad G(w) = 0 \tag{1b}$$

$$H(w) \geq 0 \tag{1c}$$

Terminology

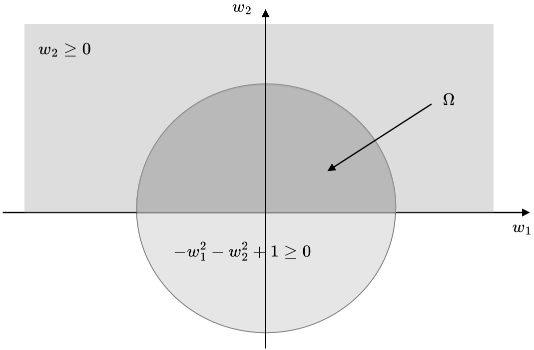
- w - decision variable
- F: objective/cost function
- G, H: equality and inequality constraint functions

- Optimization is a powerful tool used in all quantitative sciences
- Only in few special cases a closed form solution exist
- Use an iterative algorithm to find solution
- The optimization problem may be parametric, and all functions depend on a fixed parameter p

Basic definitions: the feasible set

Definition

The feasible set of the optimization problem (1) is defined as
$\Omega = \{w \in \mathbb{R}^n \mid G(w) = 0, H(w) \geq 0\}$. A point $w \in \Omega$ is is called a feasible point.



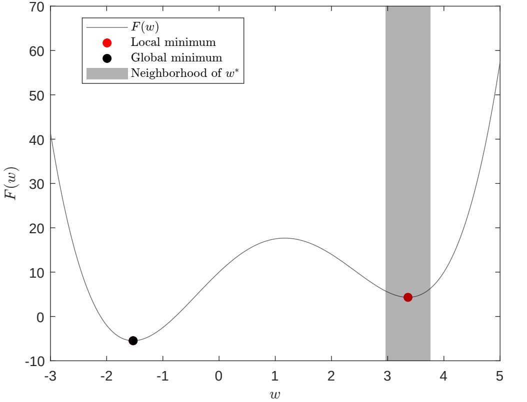
The feasible set is the intersection of the two grey areas (halfspace and circle)

Basic definitions: local and global minimizer

Definition

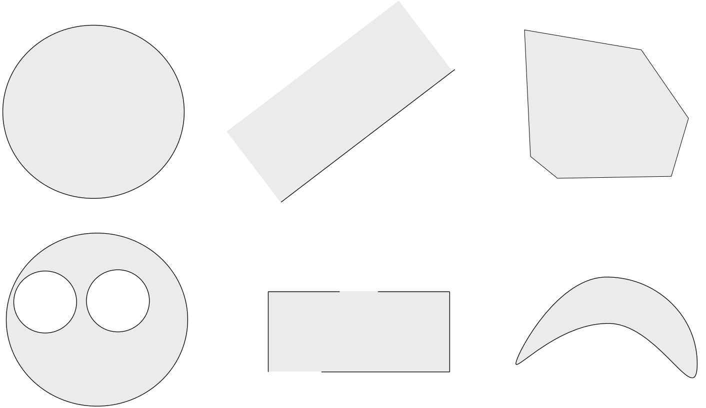
- A point $w^* \in \Omega$ is called a **local minimizer** of the NLP (1) if there exists an open ball $\mathcal{B}_\epsilon(w^*)$ with $\epsilon > 0$, such that for all $w \in \mathcal{B}_\epsilon(w^*) \cap \Omega$ it holds that $F(w) \geq F(w^*)$.
- A point $w^* \in \Omega$ is called a **global minimizer** of the NLP (1) if for all $w \in \Omega$ it holds that $F(w) \geq F(w^*)$.

The value $F(w^*)$ at a local/global minimizer w^* is called local/global minimum.



$$F(w) = \frac{1}{2}w^4 - 2w^3 - 3w^2 + 12w + 10$$

Convex sets



A set Ω is said to be convex if for any w_1, w_2 and any $\theta \in [0, 1]$ it holds $\theta w_1 + (1 - \theta) w_2 \in \Omega$

Convex functions

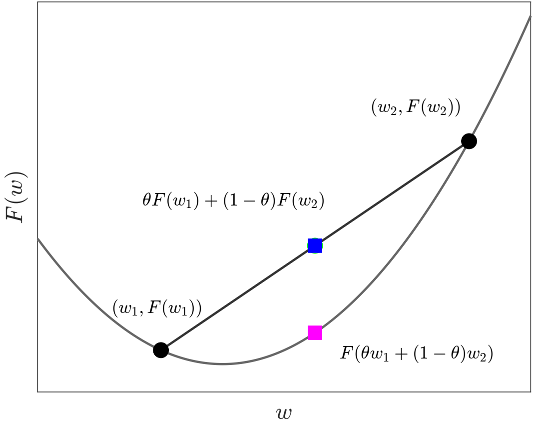
- A function F is convex if for every $w_1, w_2 \in \mathbb{R}^n$ and $\theta \in [0,1]$ it holds that

  $$F(\theta w_1 + (1-\theta) w_2) \leq \theta F(w_1) + (1-\theta) F(w_2)$$

- F is concave if and only if $-F$ is convex
- F is convex if and only if the epigraph

  $$\mathrm{epi} F = \{(w,t) \in \mathbb{R}^{n_w+1} \mid F(w) \leq t\}$$

  is a convex set



Convex optimization problems

A convex optimization problem

$$\begin{aligned} \min_{w} \quad & F(w) \\ \text{s.t.} \quad & G(w) = 0 \\ & H(w) \geq 0 \end{aligned}$$

An optimization problem is convex if the objective function F is convex and the feasible set Ω is convex.

- Every locally optimal solution is global
- First order conditions are necessary and sufficient (we come back to this)
- *"...in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."* R. T. Rockafellar, SIAM Review, 1993

Outline of the lecture

Basic definitions

Some classifications of optimization problems

Optimality conditions for unconstrained optimization

Newton's method for unconstrained optimization

Some classifications of optimization problems

Optimization problems can be:

- unconstrained ($\Omega = \mathbb{R}^n$) or constrained ($\Omega \subset \mathbb{R}^n$)
- convex or nonconvex
- linear or nonlinear
- differentiable or nonsmooth
- continuous or (mixed-)integer
- finite or infinite dimensional

Some classifications of optimization problems

Optimization problems can be:

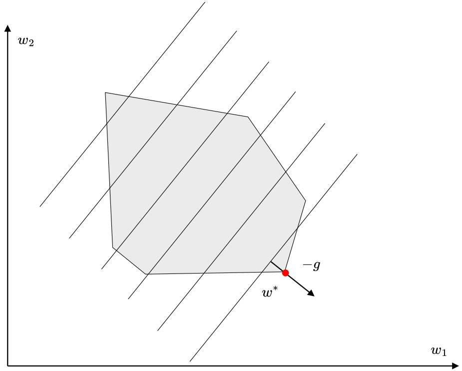
- unconstrained ($\Omega = \mathbb{R}^n$) or constrained ($\Omega \subset \mathbb{R}^n$)
- convex or nonconvex
- linear or nonlinear
- differentiable or nonsmooth
- continuous or (mixed-)integer
- finite or infinite dimensional

"... the main fact, which should be known to any person dealing with optimization models, is that in general, optimization problems are unsolvable."
Yurii Nesterov, Lectures on Convex Optimization, 2018.

Class 1: Linear Programming (LP)

Linear program

$$\min_{w} \; g^\top w$$
$$\text{s.t.} \quad Aw - b = 0$$
$$\qquad Cw - d \geq 0$$

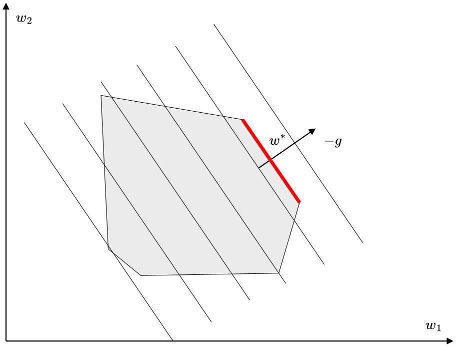


- Convex optimization problem
- 1947: simplex method by Dantzig, 1984: polynomial time interior-point method by Karmarkar
- If not unbounded, the solution is always at edge or vertex of the feasible set
- Today very mature and reliable

Class 1: Linear Programming (LP)

Linear program

$$\begin{aligned} \min_{w} \quad & g^\top w \\ \text{s.t.} \quad & Aw - b = 0 \\ & Cw - d \geq 0 \end{aligned}$$

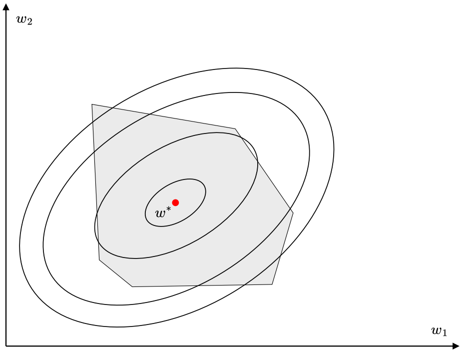


- Convex optimization problem
- 1947: simplex method by Dantzig, 1984: polynomial time interior-point method by Karmarkar
- If not unbounded, the solution is always at edge or vertex of the feasible set
- Today very mature and reliable

Class 2: Quadratic Programming (QP)

Quadratic program

$$\min_{w} \frac{1}{2} w^\top Q w + g^\top w$$
$$\text{s.t.} \quad Aw - b = 0$$
$$Cw - d \geq 0$$

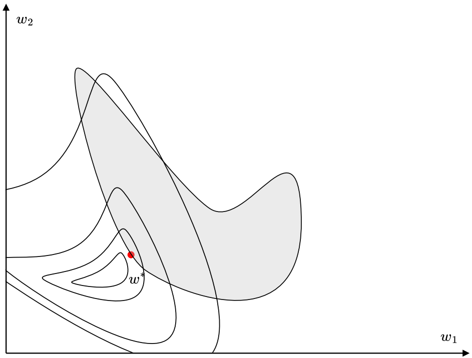


- Depending on Q, can be convex and nonconvex
- Solved online in linear model predictive control
- Many good solvers: `Gurobi`, `OSQP`, `HPIPM`, `qpOASES`, `OOQP`, ...
- Subsproblems in nonlinear optimization

Class 3: Nonlinear Program (NLP)

Nonlinear programming problem

$$\begin{aligned} \min_{w} \; & F(w) \\ \text{s.t.} \quad & G(w) = 0 \\ & H(w) \geq 0 \end{aligned}$$



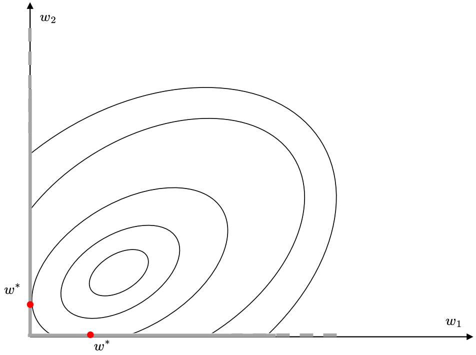
- Can be convex and nonconvex
- Solved with iterative Newton-type algorithms
- Solved in nonlinear model predictive control

Class 4: Mathematical Programs with Complementarity Constraints

MPCC

$$\begin{aligned} \min_{w_0, w_1, w_2} \quad & F(w) \\ \text{s.t.} \quad & G(w) = 0 \\ & H(w) \geq 0 \\ & 0 \leq w_1 \perp w_2 \geq 0 \end{aligned}$$

$$w = [w_0^\top, w_1^\top, w_2^\top]^\top, \; w_1 \perp w_2 \Leftrightarrow w_1^\top w_2 = 0$$

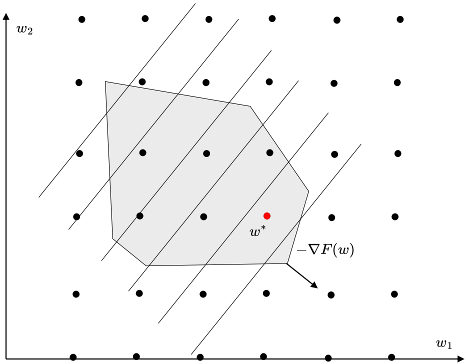


- More difficult than standard nonlinear programming since standard constraint qualifications fail to holds
- Feasible set is inherently nonsmooth and nonconvex
- Powerful modeling concept
- Requires specialized theory and algorithms

MINLP

$$\begin{aligned}\min_{w_0\in\mathbb{R}^p,\, w_1\in\mathbb{Z}^q} \quad & F(w)\\ \text{s.t.} \quad & G(w)=0\\ & H(w)\geq 0\end{aligned}$$

$w=[w_0^\top, w_1^\top]^\top, n=p+q$



- Combinatorial problem, feasible set is finite
- Branch and bound, brunch and cut methods
- Requires solution of many relaxed continuous convex or nonconvex problems

Class 6: Continuous-time Optimal Control Problems

Continuous-time Optimal Control Problem

$$\min_{x(\cdot),u(\cdot)} \int_0^T L_c(x(t),u(t))\,\mathrm{d}t + E(x(T))$$

$$\text{s.t.} \quad x(0) = \bar{x}_0$$

$$\dot{x}(t) = f_c(x(t),u(t))$$

$$0 \geq h(x(t),u(t)),\ t \in [0,T]$$

$$0 \geq r(x(T))$$

- Decision variables $x(\cdot)$ and $u(\cdot)$ in infinite dimensional function space
- Infinitely many constraints ($t \in [0,T]$)
- Smooth ordinary differential equation (ODE) $\dot{x}(t) = f_c(x(t),u(t))$
- More generally, dynamic models can be based on
  - Differential Algebraic Equations (DAE)
  - Partial Differential Equations (PDE)
  - Nonsmooth ODE
  - Stochastic ODE/PDE
- OCP can be convex or nonconvex
- All or some components of $u(t)$ may take integer values (mixed-integer OCP)

Direct optimal control methods formulate Nonlinear Programs (NLP)

Continuous-time OCP

$$\begin{aligned} \min_{x(\cdot),u(\cdot)} \quad & \int_0^T L_{\mathrm{c}}(x(t),u(t))\,\mathrm{d}t + E(x(T)) \\ \text{s.t.} \quad & x(0) = \bar{x}_0 \\ & \dot{x}(t) = f_{\mathrm{c}}(x(t),u(t)) \\ & 0 \geq h(x(t),u(t)),\ t \in [0,T] \\ & 0 \geq r(x(T)) \end{aligned}$$

Direct methods like direct collocation, multiple shooting. *First discretize, then optimize.*

Direct optimal control methods formulate Nonlinear Programs (NLP)

Continuous-time OCP

$$\min_{x(\cdot),u(\cdot)} \int_0^T L_c(x(t),u(t))\,\mathrm{d}t + E(x(T))$$

$$\begin{aligned} \text{s.t.}\quad x(0) &= \bar{x}_0 \\ \dot{x}(t) &= f_c(x(t),u(t)) \\ 0 &\geq h(x(t),u(t)),\ t\in[0,T] \\ 0 &\geq r(x(T)) \end{aligned}$$

Discrete-time OCP (an NLP)

$$\min_{x,u} \sum_{k=0}^{N-1} \ell(x_k,u_k) + E(x_N)$$

$$\begin{aligned} \text{s.t.}\quad x_0 &= \bar{x}_0 \\ x_{k+1} &= f(x_k,u_k) \\ 0 &\geq h(x_k,u_k),\ k=0,\dots,N-1 \\ 0 &\geq r(x_N) \end{aligned}$$

Direct methods like direct collocation, multiple shooting. *First discretize, then optimize.*

Nonlinear MPC solves Nonlinear Programs (NLP)

Discrete time NMPC Problem (an NLP)

$$
\begin{aligned}
\min_{x,u} \ & \textstyle\sum_{k=0}^{N-1} \ell(x_k, u_k) + E(x_N) \\
\text{s.t.} \quad & x_0 = \bar{x}_0 \\
& x_{k+1} = f(x_k, u_k) \\
& 0 \geq h(x_k, u_k), \ k = 0, \dots, N-1 \\
& 0 \geq r(x_N)
\end{aligned}
$$

Variables $x = (x_0, \dots, x_N)$ and $u = (u_0, \dots, u_{N-1})$ can be summarized in vector $w = (x, u) \in \mathbb{R}^n$.

Nonlinear MPC solves Nonlinear Programs (NLP)

Discrete time NMPC Problem (an NLP)

$$\begin{aligned} \min_{x,u} \; & \sum_{k=0}^{N-1} \ell(x_k, u_k) + E(x_N) \\ \text{s.t.} \quad & x_0 = \bar{x}_0 \\ & x_{k+1} = f(x_k, u_k) \\ & 0 \geq h(x_k, u_k), \; k = 0, \ldots, N-1 \\ & 0 \geq r(x_N) \end{aligned}$$

Variables $x = (x_0, \ldots, x_N)$ and $u = (u_0, \ldots, u_{N-1})$ can be summarized in vector $w = (x, u) \in \mathbb{R}^n$.

Nonlinear Program (NLP)

$$\begin{aligned} \min_{w \in \mathbb{R}^n} \; & F(w) \\ \text{s.t.} \; & G(w) = 0 \\ & H(w) \geq 0 \end{aligned}$$

Outline of the lecture

Basic definitions

Some classifications of optimization problems

Optimality conditions for unconstrained optimization

Newton's method for unconstrained optimization

Algebraic characterization of **unconstrained** local optima

Consider the unconstrained problem: $\min_{w \in \mathbb{R}^n} \quad F(w)$

First-Order **Necessary** Condition of Optimality (FONC)

$$w^* \text{ local optimum} \quad \Rightarrow \quad \nabla F(w^*) = 0, \; w^* \text{ stationary point}$$

Second-Order **Necessary** Condition of Optimality (SONC)

$$w^* \text{ local optimum} \quad \Rightarrow \quad \nabla^2 F(w^*) \succeq 0$$

Algebraic characterization of **unconstrained** local optima

Consider the unconstrained problem: $\min_{w \in \mathbb{R}^n} \quad F(w)$

First-Order **Necessary** Condition of Optimality (FONC)

$$w^* \text{ local optimum} \quad \Rightarrow \quad \nabla F(w^*) = 0, \;\; w^* \text{ stationary point}$$

Second-Order **Necessary** Condition of Optimality (SONC)

$$w^* \text{ local optimum} \quad \Rightarrow \quad \nabla^2 F(w^*) \succeq 0$$

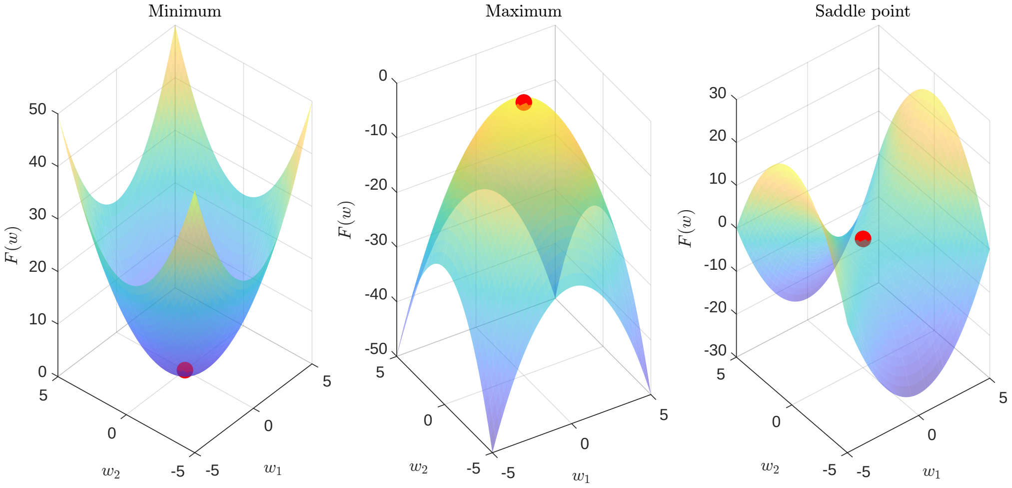
Second-Order **Sufficient** Conditions of Optimality (SOSC)

$$\nabla F(w^*) = 0 \text{ and } \nabla^2 F(w^*) \succ 0 \quad \Rightarrow \quad x^* \text{ strict local minimum}$$

$$\nabla F(w^*) = 0 \text{ and } \nabla^2 F(w^*) \prec 0 \quad \Rightarrow \quad x^* \text{ strict local maximum}$$

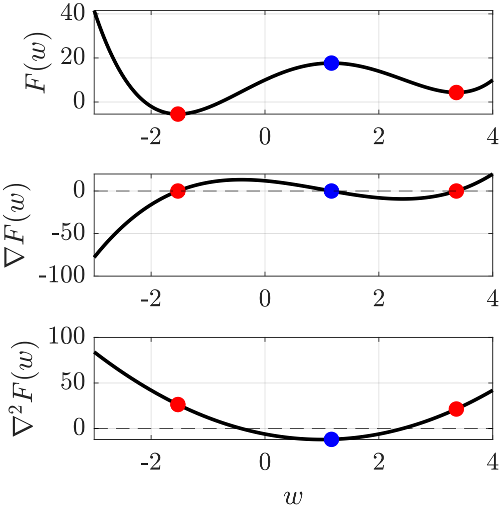
No conclusion can be drawn in the case $\nabla^2 F(w^*)$ is indefinite!

Type of stationary points



A stationary point can be a minimum, maximum or a saddle point

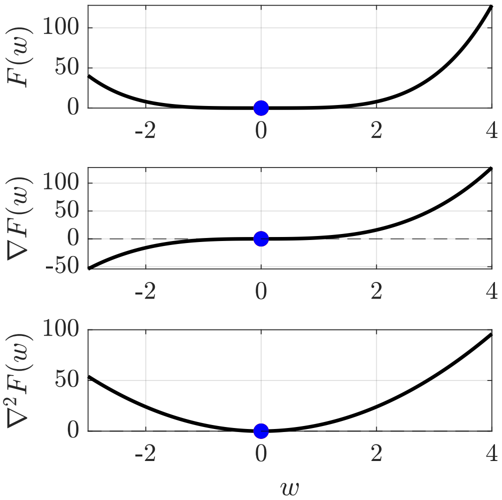
Optimality conditions - unconstrained



- Necessary conditions: find a candidate point (or to exclude points)
- Sufficient conditions: verify optimality of a candidate point

Optimality conditions - unconstrained

- Necessary conditions: find a candidate point (or to exclude points)
- Sufficient conditions: verify optimality of a candidate point
- A minimizer must satisfy SONC, but does not have to satisfy SOSC



Newton's method for unconstrained optimization

We want to solve $\min_{w \in \mathbb{R}^{n_w}} f(w)$ with $f : \mathbb{R}^{n_w} \to \mathbb{R}$ twice continuously differentiable.

Iterative Algorithm

An "iterative algorithm" generates a sequence $x_0, x_1, x_2, \ldots$ of so called "iterates" with $x_k \to x^*$.

We look for stationary points therefore we regard the equation

Newton's method for unconstrained optimization

We want to solve $\min_{w \in \mathbb{R}^{n_w}} f(w)$ with $f : \mathbb{R}^{n_w} \to \mathbb{R}$ twice continuously differentiable.

Iterative Algorithm

An "iterative algorithm" generates a sequence $x_0, x_1, x_2, \ldots$ of so called "iterates" with $x_k \to x^*$.

We look for stationary points therefore we regard the equation

$$\nabla f(w^*) = 0.$$

Idea: linearize the nonlinear equation at w_k to compute $w_{k+1} = w_k + p_k$

$$\nabla f(w_k) + \underbrace{\frac{\partial}{\partial w}(\nabla f(w_k))}_{\nabla^2 f(w_k)} p_k = 0$$

$$-(\nabla^2 f(w_k))^{-1} \nabla f(w_k) = p_k$$

p_k is called the "Newton-step", $\nabla^2 f(w_k)$ the Hessian.

An Alternative View on the Newton's Method for Optimization

The Newton's method can be obtained by a quadratic objective function, i.e. a second order Taylor approximation

Let m_k the quadratic model with objective f

$$\begin{aligned} m_k(w_k + p) &= f(w_k) + \nabla f(w_k)^\top p + \frac{1}{2} p^\top \nabla^2 f(w_k) p \\ &\approx f(w_k + p) \end{aligned}$$

We obtain p_k by

$$p_k = \arg\min_p m_k(w_k + p)$$

which translates to solving the equation

$$\nabla m(w_k + p) = \nabla f(w_k) + \nabla^2 f(w_k) p = 0$$

$$p_k = -(\nabla^2 f(w_k))^{-1} \nabla f(w_k)$$

Same formula but different interpretation!

Newton's method illustration for solving $F(w) = 0$

Linearization of F at linearization point $\bar{w}$

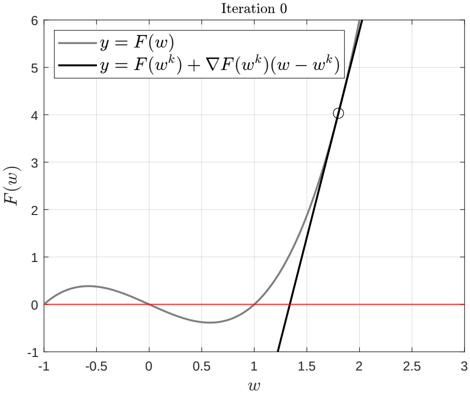
equals

First order Taylor series at $\bar{w}$

equals

$$F_{\mathrm{L}}(w; \bar{w}) := F(\bar{w}) + \frac{\partial F}{\partial w}(\bar{w}) \quad (w - \bar{w})$$

(for continuously differentiable $F : \mathbb{R}^n \to \mathbb{R}^n$)



Newton's method illustration for solving $F(w) = 0$

Linearization of F at linearization point $\bar{w}$

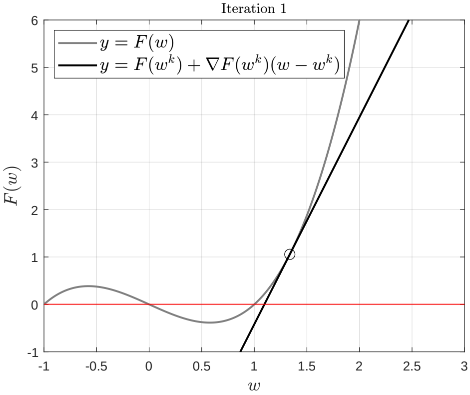
equals

First order Taylor series at $\bar{w}$

equals

$$F_{\mathrm{L}}(w; \bar{w}) := F(\bar{w}) + \nabla_w F(\bar{w})^\top (w - \bar{w})$$

(for continuously differentiable $F : \mathbb{R}^n \to \mathbb{R}^n$)



Newton's method illustration for solving $F(w) = 0$

Linearization of F at linearization point $\bar{w}$

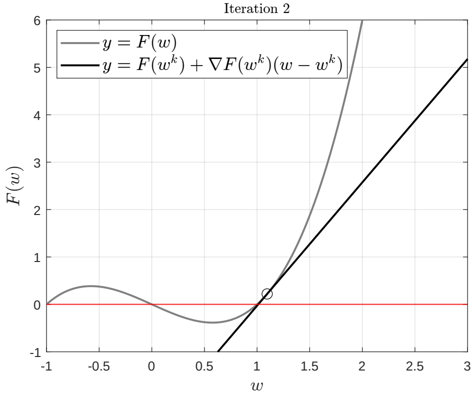
equals

First order Taylor series at $\bar{w}$

equals

$$F_{\mathrm{L}}(w; \bar{w}) := F(\bar{w}) + \nabla_w F(\bar{w})^\top (w - \bar{w})$$

(for continuously differentiable $F : \mathbb{R}^n \to \mathbb{R}^n$)



Newton's method illustration for solving $F(w) = 0$

Linearization of F at linearization point $\bar{w}$

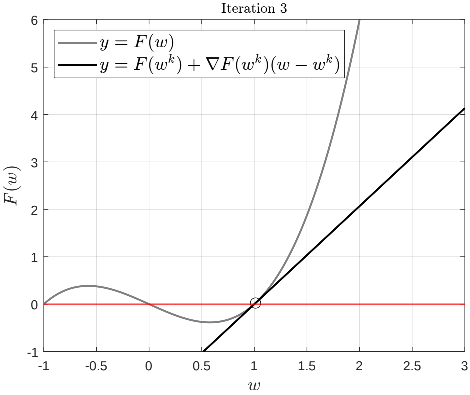
equals

First order Taylor series at $\bar{w}$

equals

$$F_{\mathrm{L}}(w; \bar{w}) := F(\bar{w}) + \nabla_w F(\bar{w})^\top (w - \bar{w})$$

(for continuously differentiable $F : \mathbb{R}^n \to \mathbb{R}^n$)



Outline of the lecture

Basic definitions

Some classifications of optimization problems

Optimality conditions for unconstrained optimization

Newton's method for unconstrained optimization

References

- Moritz Diehl, "Numerical optimization," Lecture notes, 2017.
- Jorge Nocedal, Stephen J. Wright, Numerical optimization. New York, NY: Springer New York, 2006.
- Stephen Boyd, Lieven Vandenberghe, Convex optimization. Cambridge University Press, 2004.