# Introduction to Constrained Optimization

Lecture 01A

John Bagterp Jørgensen

*Department of Applied Mathematics and Computer Science*
*Technical University of Denmark*

02612 Constrained Optimization

# Course Learning Objectives

Apply and implement numerical algorithms for constrained optimization in problems relevant to engineering.

# Outline

Problems

Applications

Parameter Estimation

Constrained Optimization

Solution of Systems of Linear Equations

# Lecture 1 - Readings

1. Nocedal & Wright: Chapter 1: Introduction, pp 1-9
2. Lecture Notes: Chapter 1: Introduction. Appendix A: Derivatives
3. Nocedal & Wright: Appendix A.2, pp 617-634
4. Nocedal & Wright: Chapter 8, Finite-Difference Derivative Approximation, pp. 193-204
5. Nocedal & Wright: Appendix A.1, pp 598-617

# Lecture 1 - Exercises

Lecture notes
Appendix A: Derivatives

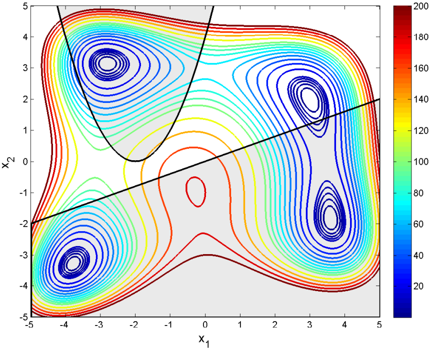
1. Problem 1: Gradient and Hessian of Multivariate Scalar Function
2. Problem 2: Rosenbrock Function
3. Problem 3: Derivatives of a Multivariate Vector Function

Lecture 01B Read about `fmincon` in the Matlab documentation and do the Himmelblau optimization problem as presented in the slides

# Constrained Optimization

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$

$$\begin{aligned} s.t. \quad & c_i(x) = 0 \qquad i \in \mathcal{E} \\ & c_i(x) \geq 0 \qquad i \in \mathcal{I} \end{aligned}$$




$$f : \mathbb{R}^n \mapsto \mathbb{R} \qquad f \in \mathcal{C}^2(\mathbb{R}^n)$$

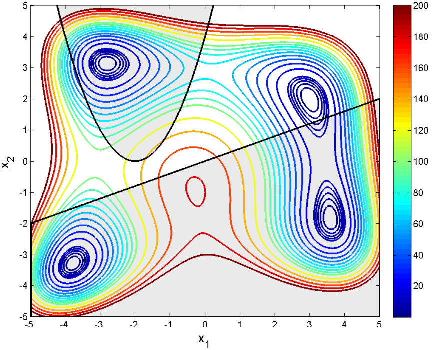
$$c_i : \mathbb{R}^n \mapsto \mathbb{R} \qquad c_i \in \mathcal{C}^2(\mathbb{R}^n)$$

# Constrained Optimization Problem

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \quad f(x_1,x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

$$s.t. \quad c_1(x_1,x_2) = (x_1+2)^2 - x_2 \geq 0$$

$$c_2(x_1,x_2) = -4x_1 + 10x_2 \geq 0$$



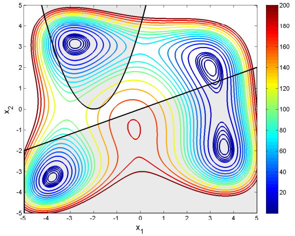
# Contour Plot with Matlab

$$
\begin{aligned}
\min_{x_1,x_2} \quad & f(x_1,x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 \\
s.t. \quad & c_1(x_1,x_2) = (x_1+2)^2 - x_2 \geq 0 \\
& c_2(x_1,x_2) = -4x_1 + 10x_2 \geq 0
\end{aligned}
$$

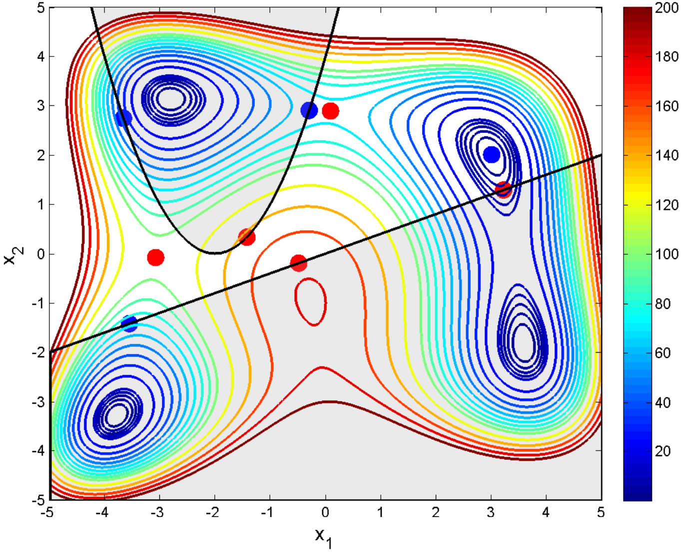
```
x = -5:0.005:5;
y = -5:0.005:5;
[X,Y] = meshgrid(x,y);
F = (X.^2+Y-11).^2 + (X + Y.^2 - 7).^2;

v = [0:2:10 10:10:100 100:20:200]
[c,h]=contour(X,Y,F,v,'linewidth',2);
colorbar

yc1 = (x+2).^2;
yc2 = (4*x)/10;
hold on
    fill(x,yc1,[0.7 0.7 0.7],'facealpha',0.2)
    fill([x x(end) x(1)],[yc2 -5 -5],[0.7 0.7 0.7],'facealpha',0.2)
hold off
```

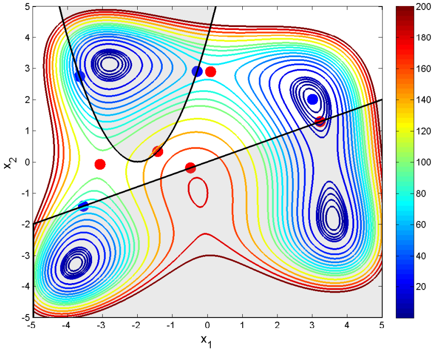


# Constrained Optimization



# Constrained Optimization

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ s.t. \quad & c_i(x) = 0 \qquad i \in \mathcal{E} \\ & c_i(x) \geq 0 \qquad i \in \mathcal{I} \end{aligned}$$




$$f : \mathbb{R}^n \mapsto \mathbb{R} \qquad f \in \mathcal{C}^2(\mathbb{R}^n)$$

$$c_i : \mathbb{R}^n \mapsto \mathbb{R} \qquad c_i \in \mathcal{C}^2(\mathbb{R}^n)$$

Feasible Region

$$\Omega = \{x \in \mathbb{R}^n : \quad c_i(x) = 0,\ i \in \mathcal{E}, \quad c_i(x) \geq 0,\ i \in \mathcal{I}\}$$

$$\min_{x\in\Omega} f(x)$$

# Convex Program

# Convex Programming

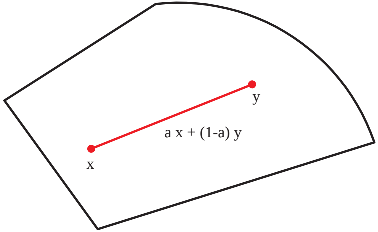
$$\min_{x \in \Omega} f(x)$$

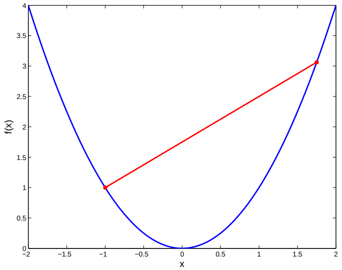
A convex set, $\mathcal{C} \subset \mathbb{R}^n$:

$$\forall x, y \in \mathcal{C}: \quad \alpha x + (1-\alpha) y \in \mathcal{C} \quad \forall \alpha \in [0,\, 1]$$

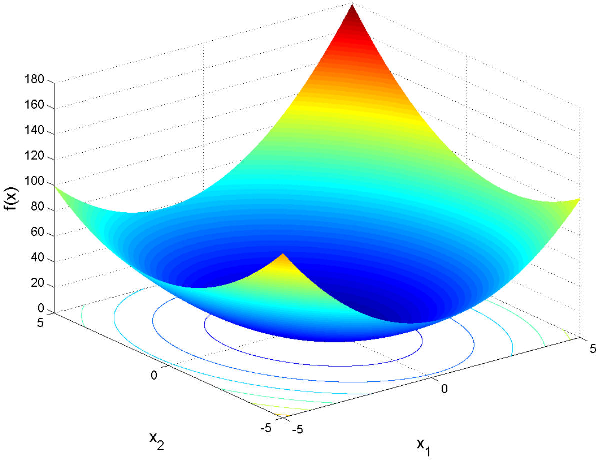
Convex function

$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$
$$\forall x, y \in \mathbb{R}^n,\ \forall \alpha \in [0,\, 1]$$

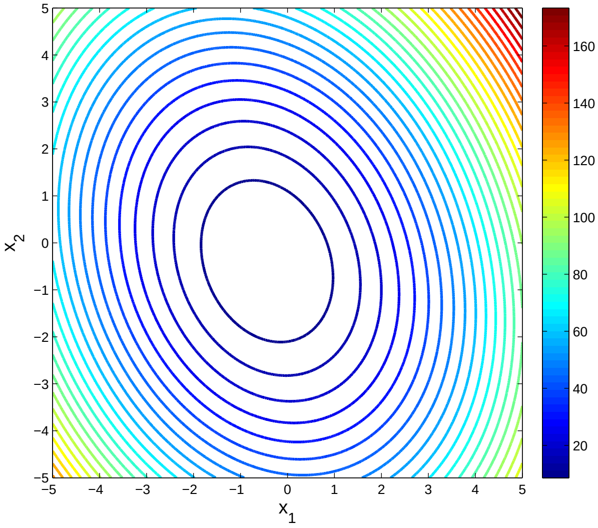


# Convex Function



# Convex Function



# Convex Program

$$\min_{x \in \mathbb{R}^n} \quad f(x) \tag{1a}$$

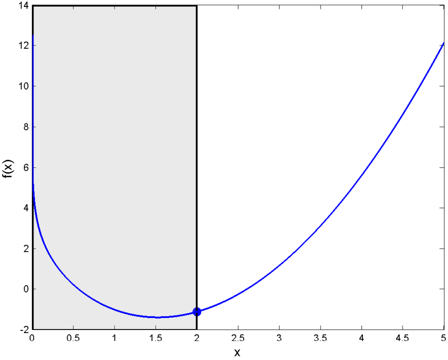
$$s.t. \quad c_i(x) = a_i' x + b_i = 0 \qquad i \in \mathcal{E} \tag{1b}$$

$$c_i(x) \geq 0 \qquad i \in \mathcal{I} \tag{1c}$$

- $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex twice continuously differentiable function
- $c_i(x)$ for $i \in \mathcal{I}$ are concave twice continuously differentiable functions.

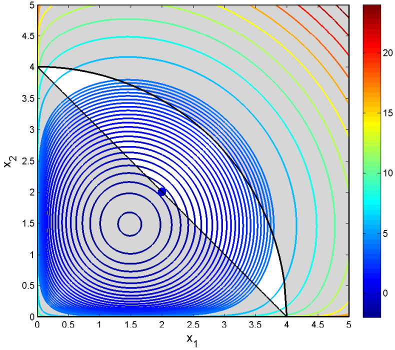
# Univariate Convex Program

$$
\begin{aligned}
\min_{x \in \mathbb{R}_{++}} \quad & f(x) = (x-1)^2 - \sqrt{x} - \ln(x) \\
s.t. \quad & c_1(x) = x - 2 \geq 0
\end{aligned}
$$



# Convex Program

$$
\begin{array}{ll}
\min\limits_{x \in \mathbb{R}^2_{++}} & f(x) = (x_1 - 1)^2 + (x_2 - 1)^2 - \sqrt{x_1 + x_2} - \ln(x_1) - \ln(x_2) \\
s.t. & c_1(x) = x_1 + x_2 - 4 \geq 0 \\
 & c_2(x) = -x_1^2 - x_2^2 + 16 \geq 0
\end{array}
$$



# Convex Programming

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) && && \text{(2a)} \\
s.t. \quad & c_i(x) = a_i' x + b_i = 0 && i \in \mathcal{E} && \text{(2b)} \\
& c_i(x) \geq 0 && i \in \mathcal{I} && \text{(2c)}
\end{aligned}
$$

- $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex function
- Equality constraints are affine (linear) functions: $c_i(x) = a_i' x + b_i = 0,\ i \in \mathcal{E}$
- Inequality constraints: $c_i(x)$ are concave functions for $i \in \mathcal{I}$. ($-c_i(x) \leq 0$, $-c_i(x)$ are convex functions for $i \in \mathcal{I}$)

# Convex Programming Problems

$$\min_{x \in \mathbb{R}^n} \quad f(x) \tag{3a}$$

$$s.t. \quad c_i(x) = a_i' x + b_i = 0 \qquad i \in \mathcal{E} \tag{3b}$$

$$c_i(x) \geq 0 \qquad i \in \mathcal{I} \tag{3c}$$

- Linear program (LP):
  $f(x) = g'x + \rho$, $c_i(x) = a_i'x + b_i$ for $i \in \mathcal{I}$
- Convex quadratic program (QP);
  $f(x) = \frac{1}{2}x'Hx + g'x + \rho$, $c_i(x) = a_i'x + b_i$ for $i \in \mathcal{I}$
- Second-Order Cone Program (SOCP)
- Semi-Definite Program (SDP)

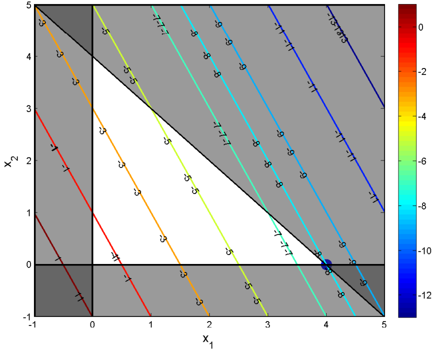
# Linear Program

# Linear Program

$$\min_{x \in \mathbb{R}^n} \quad f(x) = g'x + \rho \tag{4a}$$

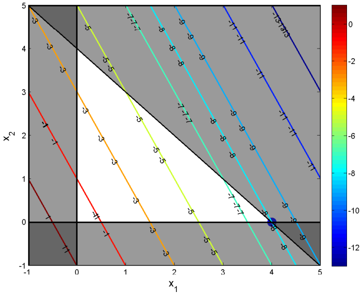
$$s.t. \quad c_i(x) = a_i'x + b_i = 0 \qquad i \in \mathcal{E} \tag{4b}$$

$$c_i(x) = a_i'x + b_i \geq 0 \qquad i \in \mathcal{I} \tag{4c}$$

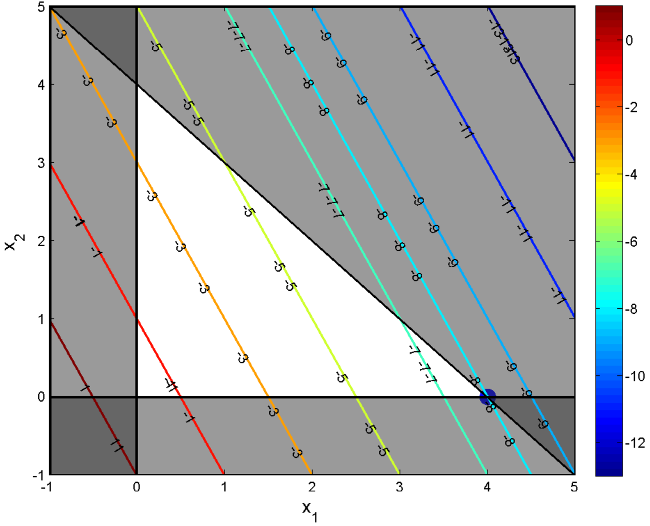


# Linear Program

$$
\begin{aligned}
\min_{x \in \mathbb{R}^2} \quad & f(x) = -2x_1 - x_2 \\
s.t. \quad & c_1(x) = x_1 \geq 0 \\
& c_2(x) = x_2 \geq 0 \\
& c_3(x) = -x_1 - x_2 + 4 \geq 0
\end{aligned}
$$



# Linear Program



# Convex Quadratic Program

# Convex Quadratic Program

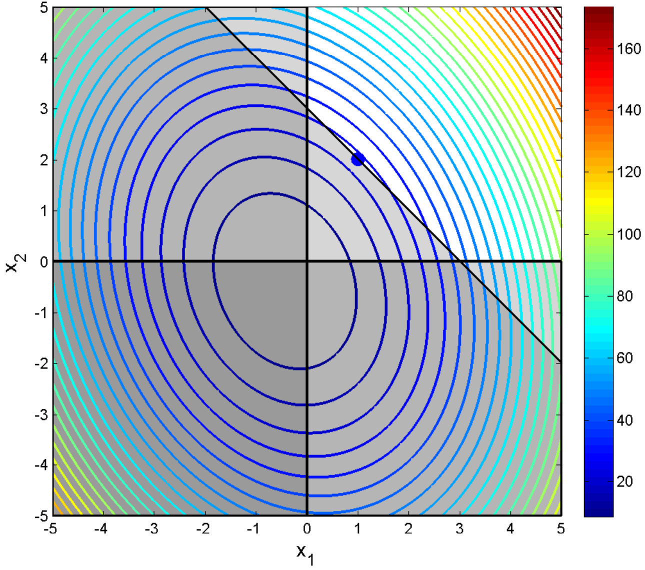
$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) = \frac{1}{2}x'Hx + g'x + \rho && \text{(5a)}\\
s.t. \quad & c_i(x) = a_i'x + b_i = 0 \qquad i \in \mathcal{E} && \text{(5b)}\\
& c_i(x) = a_i'x + b_i \geq 0 \qquad i \in \mathcal{I} && \text{(5c)}
\end{aligned}
$$

$$H \succ 0$$

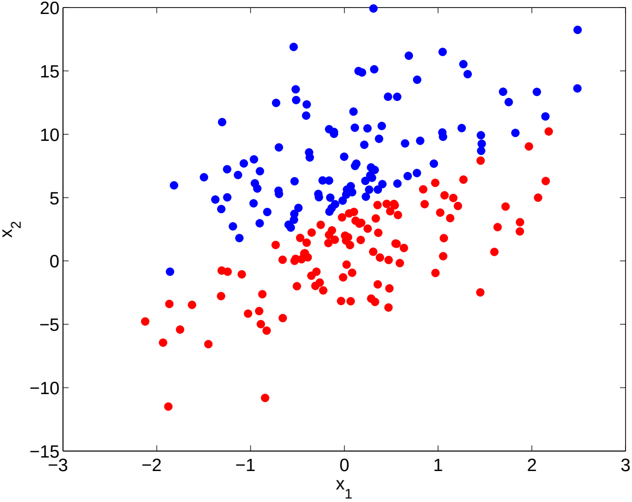
# Convex Quadratic Program

$$\begin{aligned}
\min_{x\in\mathbb{R}^2} \quad & f(x) = 3x_1^2 + 2x_2^2 + x_1x_2 + 3x_1 + 2x_2 + 4 \\
s.t. \quad & c_1(x) = x_1 \geq 0 \\
& c_2(x) = x_2 \geq 0 \\
& c_3(x) = x_1 + x_2 - 3 \geq 0
\end{aligned}$$

# Convex Quadratic Program

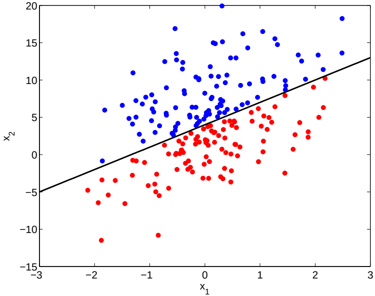


# Classification and Support Vector Machines (SVM)




Construct a classifier that can distinguish red from blue

# Classification and Support Vector Machines (SVM)




Separating hyperplane

$$3x_1 - x_2 + 4 = 0$$

The training data $\{x_k, y_k\}_{k=1}^m$ may be used to find the hyperplane as the solution of

$$\begin{array}{llll}
\min\limits_{w,b,\{\varepsilon_k\}} & w'w + c' \sum\limits_{k=1}^{m} \varepsilon_k & & \\
s.t. & w'x_k + b \geq 1 - \varepsilon_k & \text{if } y_k = 1 & k = 1, \ldots, m \\
 & w'x_k + b \leq -1 + \varepsilon_k & \text{if } y_k = -1 & k = 1, \ldots, m \\
 & \varepsilon_k \geq 0 & & k = 1, \ldots, m
\end{array}$$

This is a quadratic program

# Markowitz Portfolio Optimization Problem

Portfolio: $x \in \mathbb{R}^n$, $0 \leq x \leq 1$, $x_i$: fraction of budget invested in asset $i$

Model for return of assets: $\mathbf{y} \sim N(\mu, H)$

Portfolio return: $\mathbf{r} = \sum_{i=1}^{n} \mathbf{y}_i x_i = \mathbf{y}' x$

Expected return: $r = E\{\mathbf{r}\} = E\{\mathbf{y}'x\} = E\{\mathbf{y}\}' x = \mu' x$

Variance (risk):

$$\begin{aligned} V(x) &= E\left\{(\mathbf{r} - r)(\mathbf{r} - r)'\right\} = E\left\{(\mathbf{y}'x - \mu'x)(\mathbf{y}'x - \mu'x)'\right\} \\ &= x' E\left\{(\mathbf{y} - \mu)(\mathbf{y} - \mu)'\right\} x = x' H x \end{aligned}$$

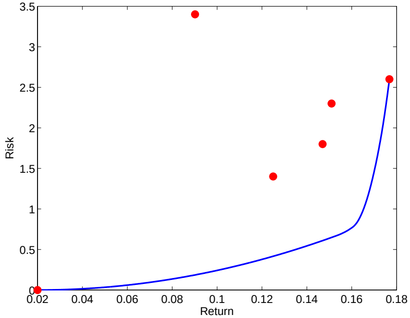
Markowitz portfolio optimization problem = Convex Quadratic Program

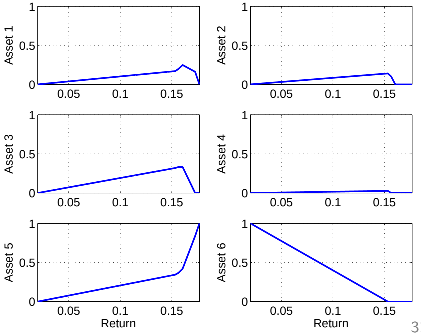
$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & V(x) = x' H x \\ s.t. \quad & \mu' x = r \\ & e' x \leq 1 \\ & x \geq 0 \end{aligned}$$

# Markowitz Portfolio Optimization Problem

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n} \quad & V(x) = x'Hx \\
s.t. \quad & \mu' x = r \\
& e'x \leq 1 \\
& x \geq 0
\end{aligned}
$$

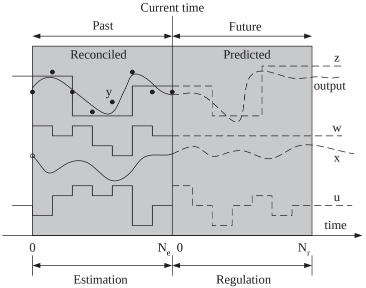
Data: $(\mu, H)$. Solve for $r \in [r_{\min}\, r_{\max}]$, to obtain the optimal portfolio and return-risk profile

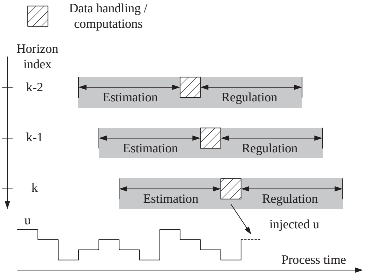


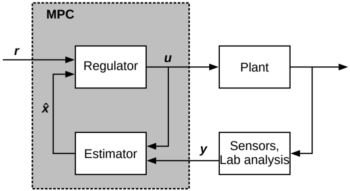



# Optimal Control
# (optimizaition of dynamical systems)

# Model Predictive Control



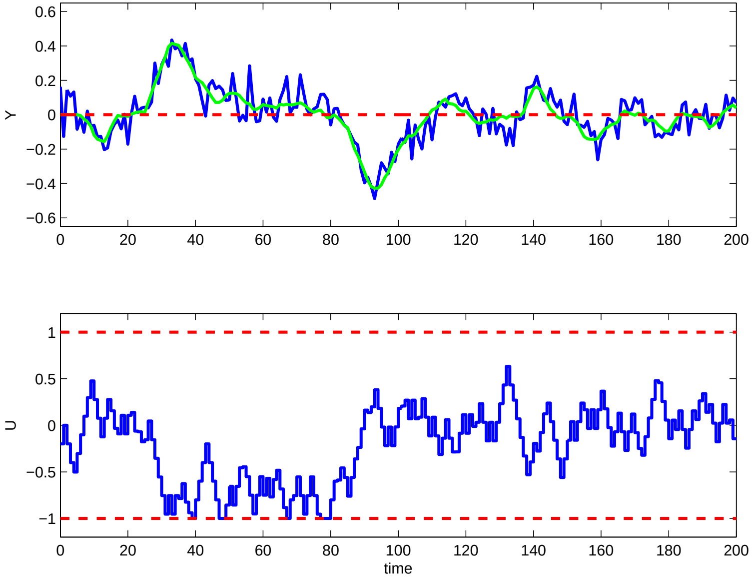




$$\min_{\{z,u\}} \quad \phi = \frac{1}{2} \sum_{k=0}^{N-1} \| z_{k+1} - r_{k+1} \|_{Q_z}^2 + \| \Delta u_k \|_S^2$$

$$\begin{aligned} s.t. \qquad & z_k = b_k + \sum_{i=1}^{n} H_i u_{k-i} && k = 1, \ldots, N \\ & u_{\min} \leq u_k \leq u_{\max} && k = 0, \ldots, N-1 \\ & \Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max} && k = 0, \ldots, N-1 \end{aligned}$$

# Model Predictive Control - Closed-Loop Profile



# Optimal Control

$$
\begin{aligned}
\min_{\{u(t),x(t)\}} \quad & J = \int_{t_0}^{t_f} g(x(t), u(t))dt + h(x(t_f)) && \text{(6a)} \\
s.t. \quad & x(t_0) = x_0 && \text{(6b)} \\
& \frac{dx}{dt} = f(x(t), u(t)) \quad t \in [t_0, t_f] && \text{(6c)}
\end{aligned}
$$

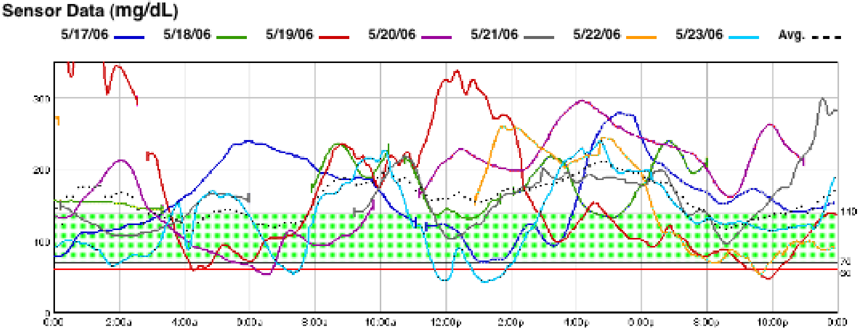
# Discrete-Time Optimal Control

$$
\begin{aligned}
\min_{\{x_{k+1},u_k\}} \quad & J = \sum_{k=0}^{N-1} g(x_k, u_k) + h(x_N) && && \text{(7a)} \\
s.t. \quad & x_0 = a && && \text{(7b)} \\
& x_{k+1} = f(x_k, u_k) && k = 0, 1, \ldots, N-1 && \text{(7c)} \\
& c(x_k, u_k) \geq 0 && k = 0, 1, \ldots, N-1 && \text{(7d)} \\
& d(x_N) \geq 0 && && \text{(7e)}
\end{aligned}
$$

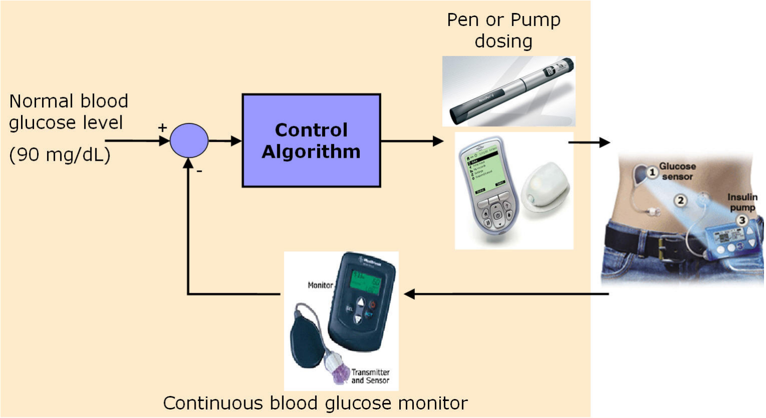
# Glucose concentration regulation

Inject insulin to keep the glucose concentration in the range 60 - 140 mg/dL.

Typical blood glucose for people with diabetes



# The artificial pancreas

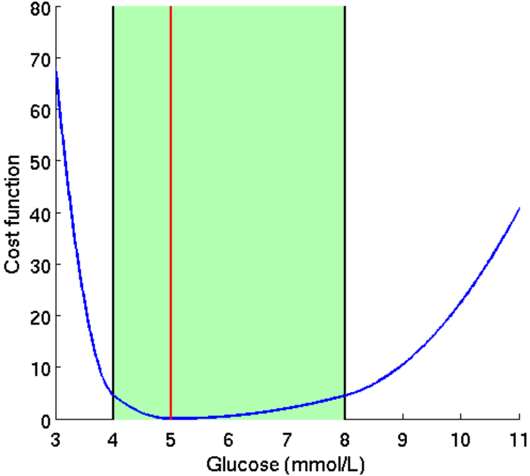


# NMPC problem formulation

$$\min_{\{u_k\}_{k=0}^{N-1}} \quad \phi = \frac{1}{2} \sum_{k=0}^{N-1} \left[ \int_{t_k}^{t_{k+1}} \kappa_1 |\max\{0, G(t) - \bar{G}\}|^2 + \kappa_2 |\max\{0, \bar{G} - G(t)\}|^2 \right.$$
$$\left. + \kappa_3 |\max\{0, G(t) - G_U\}|^2 + \kappa_4 |\max\{0, G_L - G(t)\}|^2 \right] dt$$

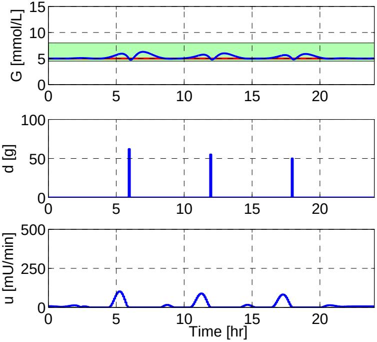
$$\begin{aligned} s.t. \qquad & x(t_0) = x_0 \\ & \dot{x}(t) = f(x(t), u(t), d(t)) \\ & y(t) = g(x(t)) \\ & u(t) = u_k \qquad t_k \leq t < t_{k+1} \\ \\ & u_{\min} \leq u_k \leq u_{\max} \\ & \Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max} \end{aligned}$$

# The objective function

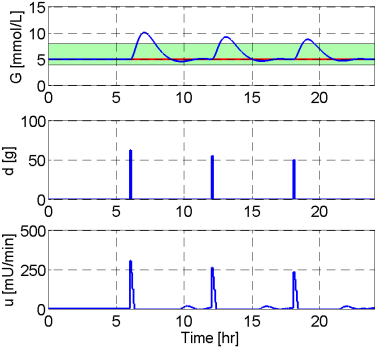


# Optimal blood glucose profile - Full state information

**Meals announced in advance.**




**Meals announced at mealtimes.**



# Parameter Estimation

# Parameter Estimation

$$\begin{aligned} \min_{x} \quad & f(x) \\ s.t. \quad & x_{\min} \le x \le x_{\max} \end{aligned}$$

- Model / prediction: $\hat{y}(x)$
- Measurement: $y$
- Error (residual): $e = e(x) = \hat{y}(x) - y$
- Covariance of error (residual): $R = R(x)$
- Objective function: $f(x)$
  - Least Squares (LS)

    $$f(x) = \frac{1}{2} \|e(x)\|_2^2$$

  - Maximum Likelihood (ML) [negative log likelihood function]

    $$f(x) = \frac{1}{2} \ln\left[\det R(x)\right] + \frac{1}{2} e(x)' R(x)^{-1} e(x)$$

# Modern Convex Optimization for Regression

$$\min_{x} \quad \phi = \sum_{k=1}^{N} \rho(e_k)$$

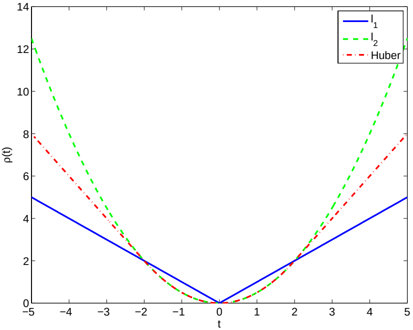
$$s.t. \quad e_k = A_k x - b_k \qquad k = 1, 2, \ldots, N$$

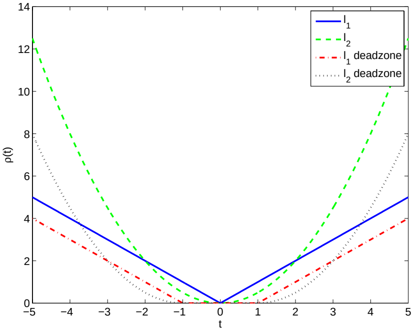
$l_1$ $\quad \rho(t) = \|t\|_1 \quad \rho(t) = \begin{cases} 0 & |t| \le \gamma \\ |t| - \gamma & |t| > \gamma \end{cases}$

$l_2$ $\quad \rho(t) = \frac{1}{2}\|t\|_2^2 \quad \rho(t) = \begin{cases} 0 & |t| \le \gamma \\ \frac{1}{2}(|t| - \gamma)^2 & |t| > \gamma \end{cases}$

Huber $\quad \rho(t) = \begin{cases} \frac{1}{2}t^2 & |t| \le \gamma \\ \gamma|t| - \frac{1}{2}\gamma^2 & |t| > \gamma \end{cases}$

SOCP, SDP



# Constrained Optimization & Numerical Linear Algebra

# Constrained Optimization Problem

Constrained optimization problem

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & f(x) && && \text{(8a)} \\
s.t. \quad & c_i(x) = 0 && i \in \mathcal{E} = \{1, 2, \dots, r\} && \text{(8b)} \\
& c_i(x) \geq 0 && i \in \mathcal{I} = \{r+1, \dots, m\} && \text{(8c)}
\end{aligned}$$

Functions

$$\begin{aligned}
& f : \mathbb{R}^n \mapsto \mathbb{R} && f \in \mathcal{C}^2 \\
& c_i : \mathbb{R}^n \mapsto \mathbb{R} && c_i \in \mathcal{C}^2 \quad i \in \mathcal{E} \cup \mathcal{I} = \{1, 2, \dots, m\}
\end{aligned}$$

# Linear System of Equations

Standard linear system of equations

$$Ax = b$$

- LU-factorization: $A$ indefinite, unsymmetric.
- Cholesky factorization: $A$ positive definite, symmetric.
- LDL-factorization: $A$ indefinite, symmetric
- QR-factorization: Often used for least-squares problems
- SVD-factorization: Rank revealing factorization

The KKT system appears very often in constrained optimization

$$\begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} g \\ b \end{bmatrix}$$

- $H$ is symmetric and positive semi-definite
- $K = \begin{bmatrix} H & -A \\ -A' & 0 \end{bmatrix}$ is non-singular
- $K$ is symmetric and indefinite

# LU Factorization

$$Ax = b$$

LU factorization with pivoting

$$PA = LU$$

- $P$ is a pivot matrix that interchanges the rows of $A$.
- $L$ is lower triangular matrix
- $U$ is an upper triangular matrix

Matlab implementation

```
[L,U,p]=lu(A,'vector');
```

Back substitutions

$$LUx = PAx = Pb$$

1. Compute: $\bar{b} = Pb$
2. Solve: $Ly = \bar{b}$
3. Solve: $Ux = y$

Matlab implementation

```
x = U\(L\b(p));
```

# Cholesky Factorization

$$Ax = b$$

$A$ is a symmetric positive definite matrix.
Positive definite matrix: $x'Ax > 0 \, \forall x \neq 0$ (all eigenvalues positive)

Cholesky factorization

$$PAP' = LL'$$

$L$ is a lower triangular matrix
$P$ is a permutation matrix

Matlab implementation

```
[L,p,s] = chol(A,'lower','vector');
```

Back-substitution
$LL'x = PAP'x = Pb$

1. Compute $\bar{b} = Pb$
2. Solve $Ly = \bar{b}$
3. Solve $L'z = y$
4. Compute $x = Pz$

Matlab implementation

```
x(s) = L'\(L\b(s));
```

# LDL factorization

$$Ax = b$$

$A$ is a symmetric indefinite matrix.

LDL factorization

$$PAP' = LDL'$$

$L$ is a lower triangular matrix
$D$ is a block-diagonal matrix
$P$ is a permutation matrix

Matlab implementation

```
[L,D,p] = ldl(A,'lower','vector');
```

Back-substitution
$LDL'x = PAP'x = Pb$

1. Compute $\bar{b} = Pb$
2. Solve $Ly = \bar{b}$
3. Solve $Dv = y$
4. Solve $L'z = v$
5. Compute $x = Pz$

Matlab implementation

```
x(p) = L'\( D \(L\b(p)));
```