Fundamentals

Eduardo Camponogara

Department of Automation and Systems Engineering
Federal University of Santa Catarina

October 2016

Summary

Optimization

Definition

Discipline concerned with:

- the computation of values for decision variables
  that induce optimal performance and
  which satisfy constraints of a mathematical model.

Problem Modeling

Need of representing the real-world:

- impossibility to interact directly with the world,
- economic hurdles,
- complexity.

We seek a representation of the world by means of well-structured models that are representative of the reality.

Problem Modeling

Need of representing the real-world:

- impossibility to interact directly with the world,
- economic hurdles,
- complexity.

We seek a representation of the world by means of well-structured models that are representative of the reality.

Problem Modeling

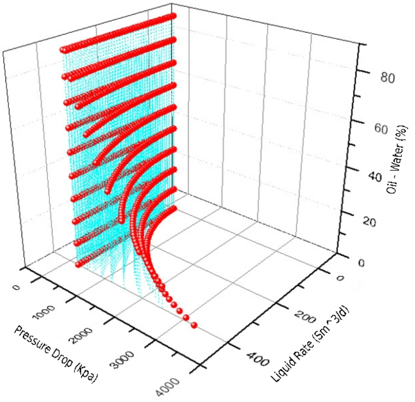
Models are simplified representations of the reality that, under some conditions and situations, are sufficiently representative.

Model Characteristics

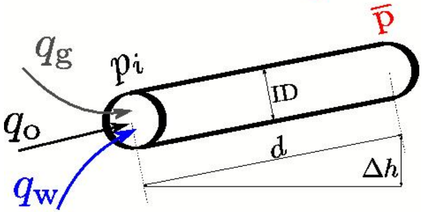
Desired characteristics:

- representation of the world;
- simplicity.

Modeling: Pressure Drop in Flowlines

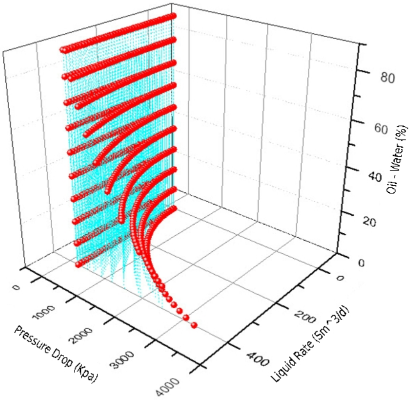


$$\Delta P = p_i - \overline{\mathrm{p}}$$

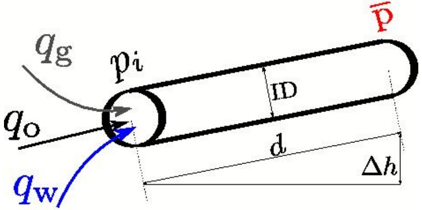


$$(q_g, q_o, q_w) \leftrightarrow (GOR, WCUT, QL)$$

Modeling: Pressure Drop in Flowlines



$$\Delta P = p_i - \overline{\mathrm{p}}$$



$$(q_g, q_o, q_w) \leftrightarrow (GOR, WCUT, QL)$$

Optimization Models

Optimization Models:

- Aim to maximize a performance criterion, such oil and gas production,
- subject to constraints that define the operating envelope.

Optimization problems are defined in "Mathematical Programming," a declarative language which is universally adopted.

Optimization Models

Optimization Models:

- Aim to maximize a performance criterion, such oil and gas production,
- subject to constraints that define the operating envelope.

Optimization problems are defined in "Mathematical Programming," a declarative language which is universally adopted.

Elements of an Optimization Formulation

1) Decision variables
2) Objective function
3) Constraints

General Formulation

$$
\begin{array}{ll}
\text{Minimize} & f(x) \\
\text{Subject to :} & \\
& g(x) \leqslant 0 \\
& h(x) = 0 \\
& x \in \mathbb{R}^n
\end{array}
$$

Summary

Linear Combination

- Let $X = \{x_1, \dots, x_n\}$ be a set of vectors in $\mathbb{R}^n$;
- A vector x is said to be a linear combination of X if:

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$
$$\alpha_1, \dots, \alpha_n \in \mathbb{R}$$

Affine Combination

- Let $X = \{x_1, \ldots, x_n\}$ be a set of vectors in $\mathbb{R}^n$;
- A vector x is said to be an affine combination of X if:

$$
\begin{gathered}
x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \\
\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1 \\
\alpha_1, \ldots, \alpha_n \in \mathbb{R}
\end{gathered}
$$

Convex Combination

- Let $X = \{x_1, \ldots, x_n\}$ be a set of vectors in $\mathbb{R}^n$;
- A vector x is said to be a convex combination of X if:

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$
$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$$
$$\alpha_1, \ldots, \alpha_n \geq 0$$

Vector Space

A set of vectors S is a vector space if:

- closed with respect to addition:

$$x + y \in S, \ \forall x, y \in S$$

- closed with respect to scalar multiplication:

$$\alpha x \in S, \ \forall x \in S, \ \forall \alpha \in \mathbb{R}$$

Affine Hull

Affine Hull of a set of vectors $X \subset \mathbb{R}^n$ is the set

$$\mathrm{aff}(X) = \{x : x \text{ is an affine combination of elements of } X\}$$

- Example: $\{x \in \mathbb{R}^n : Ax = b\}$ is an affine set.
- Exercise: prove this property.

Affine Hull

Affine Hull of a set of vectors $X \subset \mathbb{R}^n$ is the set

$$\mathrm{aff}(X) = \{x : x \text{ is an affine combination of elements of } X\}$$

- Example: $\{x \in \mathbb{R}^n : Ax = b\}$ is an affine set.
- Exercise: prove this property.

Linear Function

A function $f : \mathbb{R}^n \to \mathbb{R}$ is linear if:

- $f(\alpha x) = \alpha f(x)$ for all $x \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$.
- $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}^n$.

Taylor's Expansion

- Let $f : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable function.
- Then:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(\widehat{x})(x - \widehat{x})^k}{k!} + \mathcal{O}(|x - \widehat{x}|^{n+1})$$

- A polynomial is a universal approximator.

Taylor's Expansion

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function. Then:

$$f(x) = f(\widehat{x}) + \nabla f(\widehat{x})^{\mathrm{T}}(x - \widehat{x}) + \frac{1}{2}(x - \widehat{x})^{\mathrm{T}}\nabla^2 f(\widehat{x})(x - \widehat{x}) + \mathcal{O}(\|x - \widehat{x}\|^3)$$

in which:

- $\nabla f(\widehat{x})$ is the gradient; and
- $\nabla^2 f(\widehat{x})$ is the Hessian matrix.

Convex Function

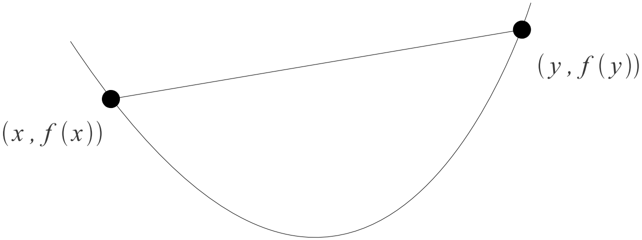
A function $f : R^n \to R$ is convex if:

- the domain of f is a convex set, and
- for all $x, y \in \mathrm{dom}\, f$, and θ such that $0 \leq \theta \leq 1$, we have that:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

Convex Function

Geometrically, the inequality means that the chord between $(x, f(x))$ and $(y, f(y))$ lies above the graph of f:



Convex Function: Example

Consider a system of linear equations:

$$Ax = b$$

If $b \notin \text{range}(A) \iff \text{rank}(A) < \text{rank}([A\ b])$, then an approximate solution can be sought by solving a least squares problem:

$$\min_x f(x) = \|Ax - b\|^2$$

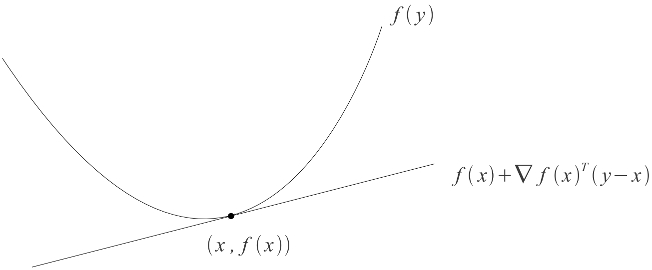
Notice that $f(x)$ is a convex function.

Convex Function: First-Order Condition

If f is differentiable, then f is convex if and only if $\mathrm{dom}\, f$ is a convex set and

$$f(y) \geq f(x) + \nabla f(x)^{\mathrm{T}}(y - x)$$

for all $x, y \in \mathrm{dom}\, f$.



Convex Function: Second-Order Condition

If f is differentiable, then f is convex if and only if $\mathrm{dom}\, f$ is a convex set and its Hessian is positive semidefinite:

$$\nabla^2 f(x) \succeq 0 \quad \forall x \in \mathrm{dom}\, f$$

Exercise: show that the least squares problem is convex.

Summary

Introduction

Fundamentals

System Identification

System Identification

- Propose a model of a system whose internal elements and connections are unknown.
- Only input signals and corresponding outputs are given.
- Carry out experiments by inputing values

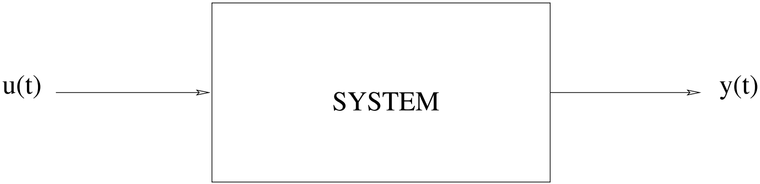
$$u(t), t = 0, \ldots, N$$

and then observing the resulting outputs

$$y(t), t = 0, \ldots, N$$

over time.

System Identification



System Identification

A simple and widely used model is the moving average:

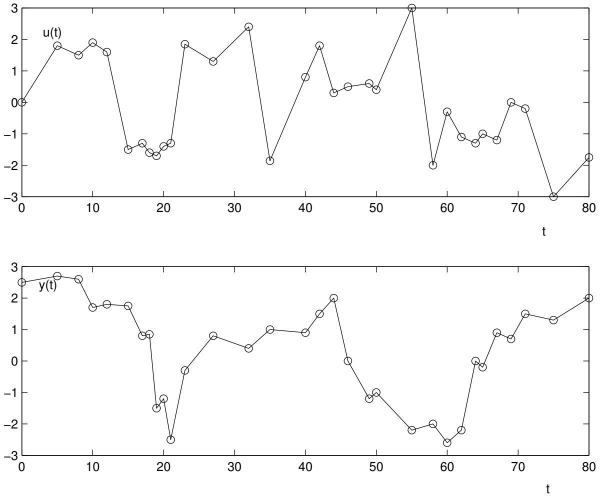
$$\hat{y}(t) = h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + \cdots + h_n u(t-n)$$

with n delays.

System Identification

- $\hat{y}(t)$ is the prediction of $y(t)$ produced by the model for the current input, $u(t)$,, and past n inputs, $u(t-1), u(t-2), \ldots, u(t-n)$.
- $h_0, h_1, \ldots, h_n$ are the parameters that define the linear combination of the inputs.

Input-Output Example



Least Squares Estimation

Find a model that minimizes the prediction error:

$$\min_{h_0, \ldots, h_n} E = \left[\sum_{t=n}^{t=N} (\hat{y}(t) - y(t))^2 \right]^{\frac{1}{2}}$$

Least Squares Estimation

Notice that the error E can be expressed in matrix form, by letting:

$$
\begin{aligned}
E &= \left[\sum_{t=n}^{t=N} (\hat{y}(t) - y(t))^2\right]^{\frac{1}{2}} \\
&= \|Ax - b\|
\end{aligned}
$$

$$
A = \begin{bmatrix}
u(n) & u(n-1) & u(n-2) & \dots & u(0) \\
u(n+1) & u(n) & u(n-1) & \dots & u(1) \\
u(n+2) & u(n+1) & u(n) & \dots & u(2) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
u(N) & u(N-1) & u(N-2) & \dots & u(N-n)
\end{bmatrix}
$$

Least Squares Estimation

$$x = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \text{ and } b = \begin{bmatrix} y_n \\ y_{n+1} \\ y_{n+2} \\ \vdots \\ y_N \end{bmatrix}$$

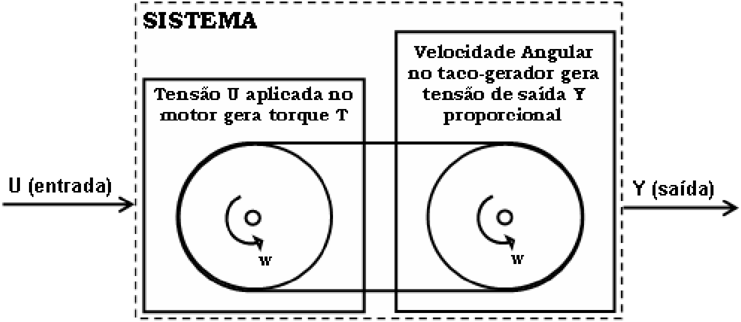
Least Squares Estimation

The estimation problem can be cast as least squares problem:

$$\min_{x \in \mathbb{R}^{n+1}} \|Ax - b\|^2$$

System Identification of a Simple Motor-Generator

- Generator connected to a motor by a belt.



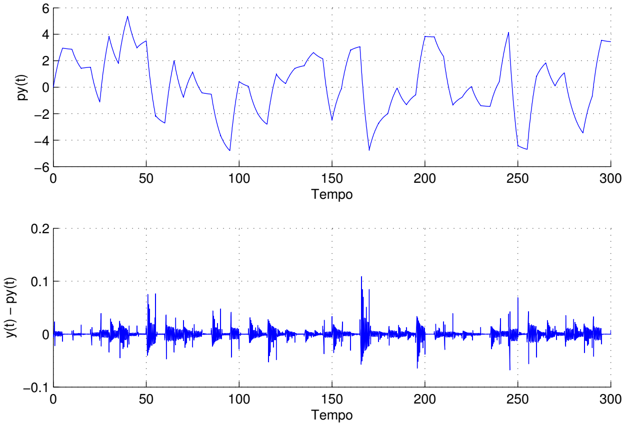
System Identification of a Simple Motor-Generator

Prediction model takes into account past outputs:

$$\hat{y}(t) = h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + \cdots + h_n u(t-n) + \\ + w_1 y(t-1) + w_2 y(t-2) + \cdots + w_n y(t-n)$$

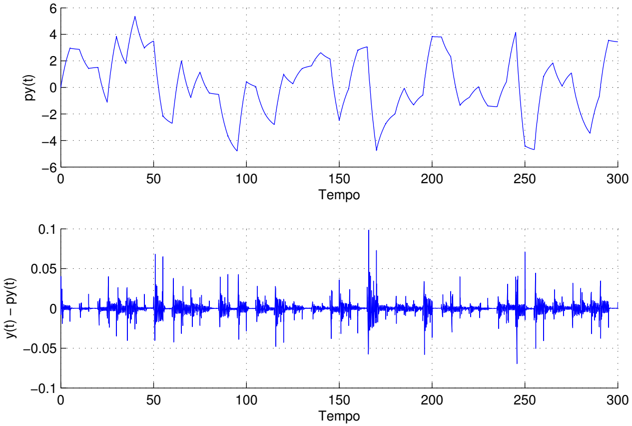
This models uses feedback, whereas a model purely based on inputs would be an open-loop model.

System Identification $(n = 3)$



The errror vector $e = y(t) - \hat{y}(t)$ has norm $\|e\| = 0.5495$.

System Identification ($n = 10$)



The error vector $e = y(t) - \hat{y}(t)$ has norm $\|e\| = 0.4778$.

Fundamentals

- Thank you for attending this lecture!!!