Parameter estimation

by Sergei Savin

Spring 2020

Content

- Parameter estimation
    - Problem statement
    - Parameter uncertainty sources
    - Problem statement - more precise
    - System Linear in Parameters
    - Example
- Least Squares in parameter estimation
    - Problem statement
    - Solution
    - Multiple measurements

Assume we have a system, whose model is defined in terms of parameters θ:

$$\dot{\mathsf{x}} = \mathsf{f}(\mathsf{x}, \mathsf{u}, \theta)$$

Task: knowing relations $\mathsf{f}(\mathsf{x}, \mathsf{u}) = \mathsf{f}(\mathsf{x}, \mathsf{u}, \theta)$, find a stabilizing control law.

In case of a linear system, the dynamics is given in the form:

$$\dot{\mathsf{x}} = \mathsf{A}(\theta)\mathsf{x} + \mathsf{B}(\theta)\mathsf{u}$$

and the task is: knowing relations $\mathsf{A} = \mathsf{A}(\theta)$ and $\mathsf{B} = \mathsf{B}(\theta)$, find a stabilizing control law.

There are a number of reasons parameters of the model can be unknown:

- Parameters that can't be measured.
- Parameters change with time.
- Unmodeled dynamics.
- Unknown forces.
- Dynamic changes in the system (picking up a load, etc).

Unfortunately, in order to treat the problem we need to assume that we can, directly or indirectly, measure higher order derivatives.

Example 1
Let us consider spring-damper system: $m\ddot{q} + \mu\dot{q} + kq = u$. Introducing notation $y = u$ we arrive at:

$$y = m\ddot{q} + \mu\dot{q} + kq$$

Example 2
Consider a general mechanical system:

$$\mathsf{H}\ddot{\mathsf{q}} + \mathsf{c}(\dot{\mathsf{q}}, \mathsf{q}) = \mathsf{u}$$

This system is *linear* with respect to parameters such as m_i, $l_i l_j$, l_i^2, $m_i l_i^2$, $m_i g$, etc., where l_i are lengths, m_i are masses, g is the gravitational acceleration. Same as before, we make $y = u$. Therefore we can rewrite it as follows:

$$\mathsf{y} = \mathsf{Y}(\ddot{\mathsf{q}}, \dot{\mathsf{q}}, \mathsf{q})\theta$$

where matrix Y can be a nonlinear function of the coordinates q and their derivatives. It is called a *regressor*.

From here on we assume that we measure y and we know that y is a function of state and it's derivatives, and parameters θ.
Moreover, we assume that y is *linear with respect to parameters* θ:

$$\mathsf{y} = \mathsf{M}(\mathsf{x}, \dot{\mathsf{x}})\theta$$

where M will be referred to as regressor matrix.

We introduce definitions:

- $\tilde{\theta}$ - estimated value of parameters θ;
- $\varepsilon_\theta = \theta - \tilde{\theta}$ - parameter estimation error;
- $\mathsf{e}_\theta = \mathsf{y} - \mathsf{M}\tilde{\theta}$ - parameter estimation output error.

Going back to the spring-damper system $y = m\ddot{q} + \mu\dot{q} + kq$, we can denote:

- $\mathsf{M} = [\ddot{q}\ \dot{q}\ q]$ - regressor matrix;
- $\theta = [m\ \mu\ k]^\top$ - parameters;

Thus we get:

$$y = [\ddot{q}\ \dot{q}\ q][m\ \mu\ k]^\top = \mathsf{M}\theta$$

Another example is a pendulum: $y = ml^2\ddot{\varphi} + mglsin(\varphi)$. We can denote:

- $\mathsf{M} = [\ddot{\varphi}\ sin(\varphi)]$ - regressor matrix;
- $\theta = [ml^2\ mgl]^\top$ - parameters;

Thus we get:

$$y = [\ddot{\varphi}\ sin(\varphi)][ml^2\ mgl]^\top = \mathsf{M}\theta$$

We want to minimize parameter estimation error ε_θ. However, we do not measure it directly. Let us instead minimize directly measured parameter estimation output error e_θ.
To this end we introduce a cost function J:

$$J = \frac{1}{2}\mathsf{e}_\theta^\top \mathsf{e}_\theta$$

Expanding the definition, we get:

$$J = \frac{1}{2}(\mathsf{y} - \mathsf{M}\tilde{\theta})^\top (\mathsf{y} - \mathsf{M}\tilde{\theta}) = \frac{1}{2}(\mathsf{y}^\top \mathsf{y} - 2\tilde{\theta}^\top \mathsf{M}^\top \mathsf{y} + \tilde{\theta}^\top \mathsf{M}^\top \mathsf{M}\tilde{\theta})$$

Derivative of J with respect to parameters estimate is:

$$\frac{\partial J}{\partial \tilde{\theta}} = -\mathsf{M}^\top \mathsf{y} + \mathsf{M}^\top \mathsf{M}\tilde{\theta}$$

We know that when the optimal estimation is found, the derivative $\frac{\partial \mathsf{J}}{\partial \tilde{\theta}}$ of the cost function will be equal to zero. Therefore:

$$\frac{\partial \mathsf{J}}{\partial \tilde{\theta}} = -\mathsf{M}^\top \mathsf{y} + \mathsf{M}^\top \mathsf{M} \tilde{\theta} = 0$$

$$\tilde{\theta} = (\mathsf{M}^\top \mathsf{M})^{-1} \mathsf{M}^\top \mathsf{y}$$

This presents the *least squares solution* for the estimation problem.

If the parameters θ are constant, we can use *multiple measurements* to find them. Let us denote the value of M matrix for i-th measurement as M_i and the corresponding value of output vector y as y_i.
Then we can introduce compound output and estimation matrices for n measurements:

$$\bar{\mathsf{M}} = \begin{bmatrix} \mathsf{M}_1 \\ ... \\ \mathsf{M}_n \end{bmatrix}$$

$$\bar{\mathsf{y}} = \begin{bmatrix} \mathsf{y}_1 \\ ... \\ \mathsf{y}_n \end{bmatrix}$$

Then we can use least squares to determine optimal parameter estimate same as before:

$$\tilde{\theta} = (\bar{\mathsf{M}}^\top \bar{\mathsf{M}})^{-1} \bar{\mathsf{M}}^\top \bar{\mathsf{y}}$$

Lecture slides are available via Moodle.

You can help improve these slides at:
`https://github.com/SergeiSa/Linear-Control-Slides-Spring-2020`

Check Moodle for additional links, videos, textbook suggestions.