XXXVII SUMMER SCHOOL ON MATHEMATICAL PHYSICS, RAVELLO, ITALY, SEPTEMBER 17 -29, 2012

#5
Building fractional-order models: Fitting experimental data using the Mittag-Leffler function

Igor Podlubny
Technical University of Kosice, Slovakia

http://www.tuke.sk/podlubny

---

826 IEEE Transactions on Dielectrics and Electrical Insulation Vol. 1 No. 5, October 1994

Capacitor Theory

Svante Westerlund
Department of Technology, University of Kalmar, Kalmar, Sweden

and Lars Ekstam
RIFA Electrolytics AB, Gränna, Sweden

The starting point in developing a new capacitor model is Curie s law of 1889 [1]. This is a purely empiric relation. A constant dc voltage U_0 applied at time $t = 0$ will produce a current

$$i(t) = \frac{U_0}{h_1 t^n} \qquad 0 < n < 1, \quad t > 0 \qquad (1)$$

h_1 is a constant related to the capacitance of the capacitor and the kind of dielectric. n is another constant, n close to 1.0 for capacitor dielectrics, and is related to the losses of the capacitor. The lower the losses, the closer to 1.0 is n. This will be verified later in this Section, see also Appendix A.

For many years we have almost daily verified Equation (1) and altogether we have measured tens of thousands of capacitors of all types and makes and have never experienced a capacitor that does not closely adhere to Equation (1). We therefore consider Equation (1) to display a normal property of all dielectrics and insulators. This is controversial! For instance von Schweidler [2,3] was of the opinion that the Curie current is abnormal and named it accordingly. Many modern workers assent to the ideas of von Schweidler, for instance [4]. But there are also a few who disagree, maybe foremost Jonscher, who in 1977 named the Curie response 'the universal dielectric response' [5].

---

826 IEEE Transactions on Dielectrics and Electrical Insulation Vol. 1 No. 5, October 1994

Capacitor Theory

Svante Westerlund
Department of Technology, University of Kalmar, Kalmar, Sweden

and Lars Ekstam
RIFA Electrolytics AB, Gränna, Sweden

Physica Scripta > Volume 43 > Number 2

Svante Westerlund 1991 *Phys. Scr.* **43** 174

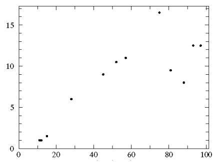
Dead matter has memory!

The contemporary art(?) of data fitting

1. Collect your experimental data

2. Which shape do they remind you?
   (How many functions do you know?
   How many combinations of those can you imagine?)

3. Use you fit in some way.
   (How can you use your fit?
   Interpolation, extrapolation – but other uses?)



"How many functions... ?"

- Linear interpolation –
  Babylonian astronomers, around 4th century BC
- Quadratic interpolation –
  Liu Zhuo, China, around 544–610
- Trigonometrical interpolation –
  Hipparchus of Rhodes, around c. 190 – 120 BC,
  Ptolemy of Alexandria, around 87–170 AD
- Approximation by circles and ellipses –
  Kepler, 1609–1619
- Approximation of growth and decay
  using exponential function –
  physics, chemistry, biology of last couple of centuries

In fact, data fitting is done with the help of solutions of differential equations

$y = kx + b$	$y'' = 0$
$y = Ce^{kx}$	$y' - ky = 0$
$y = a\sin(wx) + b\cos(wx)$	$y'' + w^2 y = 0$
$y = Ae^{kx}\sin(wx) + Be^{kx}\cos(wx)$	$a_2 y'' + a_1 y' + a_0 y = 0$

Instead of postulating the type of the fitting function, we can postulate the type of the differential equation; its coefficients must be determined.

$$Ay'' + By' + Cy = 0$$

G. M. Mittag-Leffler

Professor Donald E. Knuth,
creator of TEX:

"As far as the spacing in mathematics is concerned...
I took Acta Mathematica, from 1910 approximately; this was a journal in Sweden ... Mittag-Leffler was the editor, and his wife was very rich, and they had the highest budget for making quality mathematics printing. So the typography was especially good in Acta Mathematica."

(Questions and Answers with Prof. Donald E. Knuth, Charles University, Prague, March 1996)

nulle part elle-même, renfermant le point a, et tel que la branche de la fonction $F(x)$, formée par $\mathfrak{P}(x|a)$ et sa continuation analytique à l'intérieur de K, reste uniforme et régulière, nous désignerons cette branche par $FK(x)$.

Le problème dont nous allons nous occuper sera de trouver une représentation analytique d'une branche $FK(x)$ choisie aussi étendue que possible.

De la définition même de la fonction analytique $F(x)$, et de celle de la branche $FK(x)$, résulte immédiatement une sorte de représentation analytique de la branche $FK(x)$ en question.

En effet, pour obtenir une représentation de cette branche, il suffit d'effectuer un nombre dénombrable de prolongements analytiques de $\mathfrak{P}(x|a)$, par exemple

$$\mathfrak{P}_\nu(x|a_\nu) = \sum_{\mu=0}^{\infty} \frac{1}{\underline{|\mu}} \left(\frac{d^\mu FK(x)}{dx^\mu}\right)_{x=a_\nu} (x-a_\nu)^\mu,$$

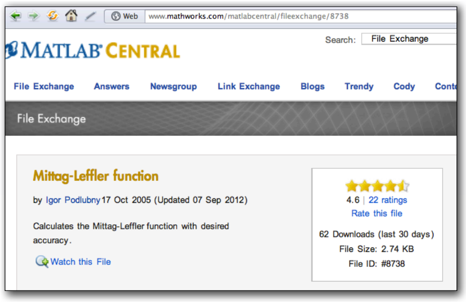
$$\nu = 0,1,2,\dots;\ a_0 = a;\ \mathfrak{P}_0(x|a) = \mathfrak{P}(x|a).$$

Les séries $\mathfrak{P}_\nu(x|a_\nu)$ sont formées au moyen des éléments

$$\left(\frac{d^\mu FK(x)}{dx^\mu}\right)_{x=a_\nu}; \qquad \left(\begin{smallmatrix}\mu=0,1,2,\dots\\ \nu=0,1,2,\dots\end{smallmatrix}\right)$$

The Mittag-Leffler function

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k+\beta)}, \quad (\alpha>0, \quad \beta>0)$$



Most used definitions of fractional differentiation

Riemann–Liouville, 1920s: (Letnikov, 1870s)

$${}_aD_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)}\left(\frac{d}{dt}\right)^n \int_a^t \frac{f(\tau)\,d\tau}{(t-\tau)^{\alpha-n+1}}, \qquad (n-1\le\alpha<n)$$

Caputo, 1967:

$${}_a^CD_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)}\int_a^t \frac{f^{(n)}(\tau)\,d\tau}{(t-\tau)^{\alpha-n+1}}, \qquad (n-1\le\alpha<n)$$

Miller–Ross, 1990s (Dzhrbashyan, 1960s)

$$D^\alpha f(t) = D^{\alpha_1}D^{\alpha_2}\dots D^{\alpha_n} f(t), \quad \left(\alpha=\sum_{k=1}^n \alpha_k;\ n-1\le\alpha<n\right)$$

Grünwald–Letnikov, 1860s (Liouville, 1830s)

$$D^\alpha f(t) = \lim_{h\to0} h^{-\alpha}\sum_{k=0}^{\left[\frac{t-a}{h}\right]}(-1)^k\binom{\alpha}{n} f(t-kh)$$

Fitting data using the Mittag-Leffler function

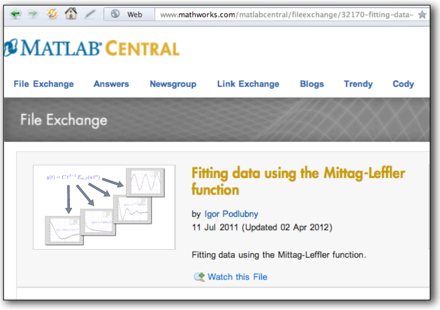
$$y = y_0\, e^{kt} \qquad y'(t) - k\,y(t) = 0, \quad y(0) = y_0$$

$$y = y_0\, t^{\beta-1} E_{\alpha,\beta}(a\,t^\alpha)$$

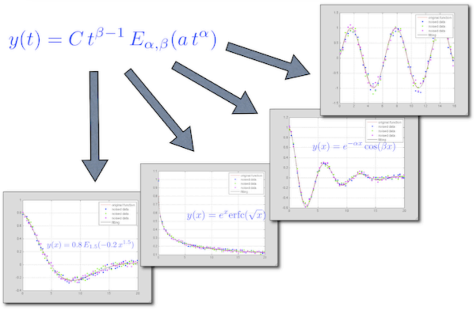
$$y = y_0\, E_{\alpha,1}(a\,t^\alpha) \qquad {}_0^CD_t^\alpha y(t) - k\,y(t) = 0, \quad y(0) = y_0$$

Fitting the experimental data with the M-L function immediately gives the basic FDE describing the process.

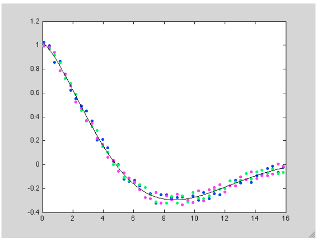
Just supply your data...



Just supply your data...



Fitting back the noised Mittag-Leffler function



Original:

$$y = y_0\, t^{\beta-1} E_{\alpha,\beta}(a\, t^{\alpha})$$

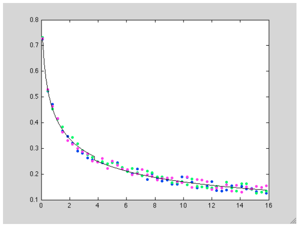
$$\alpha = 1.5, \quad \beta = 1, \quad y_0 = 1, \quad a = -0.2$$

Noise:

```
y1 = 1 * x.^(beta-1) .* mlf(alfa, beta, -0.2*x.^alfa, 7) + (-.05 + .1*rand(size(x)));
y2 = 1 * x.^(beta-1) .* mlf(alfa, beta, -0.2*x.^alfa, 7) + (-.05 + .1*rand(size(x)));
y3 = 1 * x.^(beta-1) .* mlf(alfa, beta, -0.2*x.^alfa, 7) + (-.05 + .1*rand(size(x)));
```

Fitting: $\alpha = 1.4934, \quad \beta = 0.9934 \quad y_0 = 1.0084 \quad a = -0.1985$

Fitting the complementary error function



Original:

$$y = e^t \mathrm{erfc}(\sqrt{t})$$

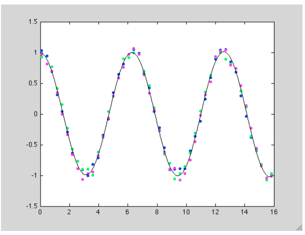
$$e^t \mathrm{erfc}(\sqrt{t}) = E_{1/2,1}(-\sqrt{t})$$

Noise:

```
y1 = exp(x).*erfc(sqrt(x))  + (-.02 + .04*rand(size(x)));
y2 = exp(x).*erfc(sqrt(x))  + (-.02 + .04*rand(size(x)));
y3 = exp(x).*erfc(sqrt(x))  + (-.02 + .04*rand(size(x)));
```

Fitting: $\alpha = 0.4503, \quad \beta = 0.9861, \quad y_0 = 1.0525, \quad a = -1.1045$

Fitting the cosine function



Original:

$$y = \cos(t)$$

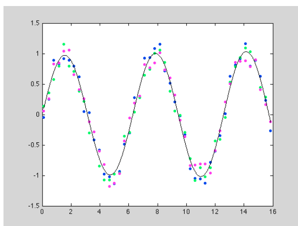
$$\cos(t) = E_{2,1}(-t^2)$$

Noise:

```
y1 = cos(x)  + (-.1 + .2*rand(size(x)));
y2 = cos(x)  + (-.1 + .2*rand(size(x)));
y3 = cos(x)  + (-.1 + .2*rand(size(x)));
```

Fitting: $\alpha = 2.0029, \quad \beta = 0.9891, \quad y_0 = 0.9855, \quad a = -0.9967$

Fitting the sine function



Original:

$$y = \sin(t)$$

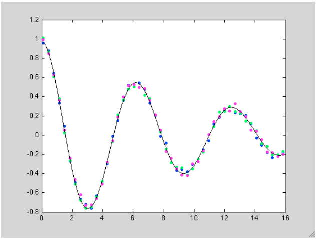
$$\sin(t) = t\, E_{2,2}(-t^2)$$

Noise:

```
y1 = sin(x)  + (-.2 + .4*rand(size(x)));
y2 = sin(x)  + (-.2 + .4*rand(size(x)));
y3 = sin(x)  + (-.2 + .4*rand(size(x)));
```

Fitting: $\alpha = 2.0056, \quad \beta = 1.9960, \quad y_0 = 0.9700, \quad a = -0.9959$

Fitting damped oscillations



Original:

$$y = e^{-0.1\,t}\,\cos(t)$$

$$y(t) = y_0 E_{\alpha,\beta}(at^{\alpha})$$

Noise:

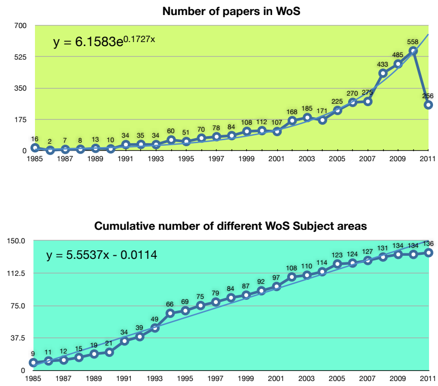
```
y1 = exp(-0.1*x).*cos(x) + (-.05 + .1*rand(size(x)));
y2 = exp(-0.1*x).*cos(x) + (-.05 + .1*rand(size(x)));
y3 = exp(-0.1*x).*cos(x) + (-.05 + .1*rand(size(x)));
```

Fitting: $\alpha = 1.8784, \quad \beta = 0.9982, \quad y_0 = 0.9731, \quad a = -1.0089$

Fitting data using the Mittag-Leffler function

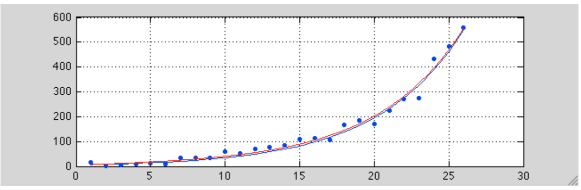
Example: Growth of the number of FC articles in Web of Science

	Number of papers in WoS	Number of WoS Subject areas in particular year	Cumulative number of different WoS Subject areas
1985	16	9	9
1986	2	4	11
1987	7	3	12
1988	8	6	15
1989	13	8	19
1990	10	9	21
1991	34	24	34
1992	35	20	39
1993	34	30	49
1994	60	38	66
1995	51	36	69
1996	70	37	75
1997	78	30	79
1998	84	43	84
1999	108	36	87
2000	112	47	92
2001	107	46	97
2002	168	58	108
2003	185	53	110
2004	171	64	114
2005	225	66	123
2006	270	57	124
2007	275	58	127
2008	433	70	131
2009	485	76	134
2010	558	61	134
2011	256	53	136



Fitting data using the Mittag-Leffler function

Example: Growth of the number of FC articles in Web of Science



$y(t) = 6.1583\,e^{0.1727t}$ $\qquad \sum (y(t_k) - y_k)^2 = 1.0586 \cdot 10^4$

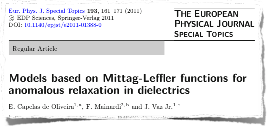
$y(t) = 4.2072\,E_{0.1559,\,2.6082}(0.7987\,t^{0.1559})$ $\qquad \sum (y(t_k) - y_k)^2 = 9.0512 \cdot 10^3$

The Queen Function

"In fact, ... , functions of Mittag-Leffler type enter as solutions of many problems dealt with fractional calculus so that they like to refer to the **Mittag-Leffler function** to as the

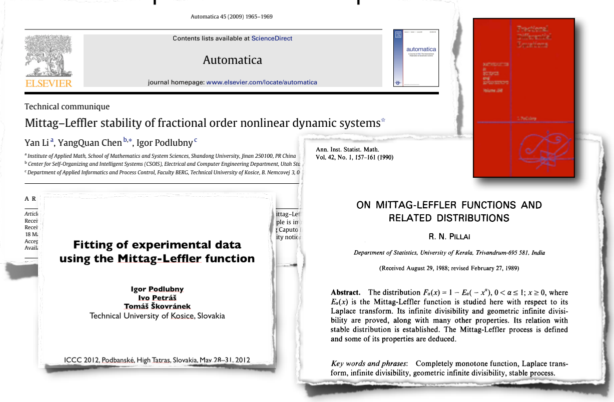
Queen function of Fractional Calculus,

in contrast with its role of a **Cinderella** function played in the past."



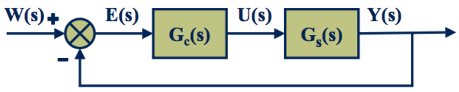
Mittag-Leffler function:

a replacement for the exponential function



From CO to VO modelling

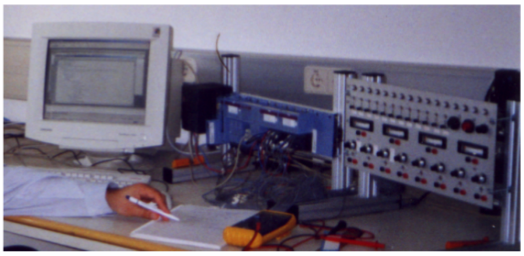
Fractional order systems and controllers



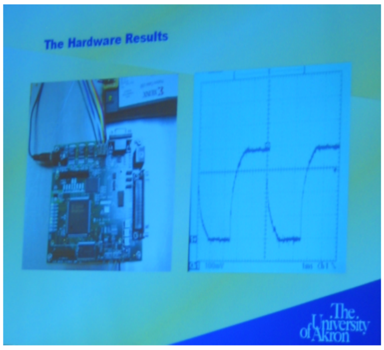
$$G_s(s)=\frac{1}{a_n s^{\beta_n}+a_{n-1}s^{\beta_{n-1}}+\ldots+a_1 s^{\beta_1}+a_0 s^{\beta_0}}$$

$$a_n D^{\beta_n}y(t)+a_{n-1}D^{\beta_{n-1}}y(t)+\ldots+a_0 D^{\beta_0}y(t)=u(t)$$

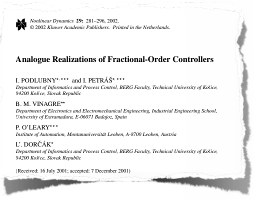
Digital realization: PLC B & R 2005



Digital realization: FPGA



Analogue realization



Nonlinear Dynamics **29:** 281–296, 2002.
© 2002 *Kluwer Academic Publishers. Printed in the Netherlands.*

Analogue Realizations of Fractional-Order Controllers

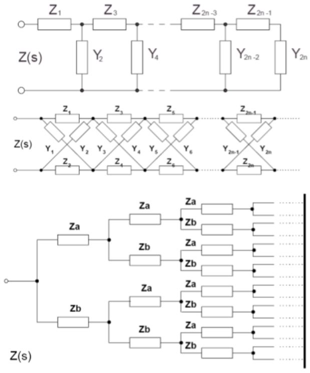
I. PODLUBNY*,*** and I. PETRÁŠ*,***
Department of Informatics and Process Control, BERG Faculty, Technical University of Košice, 04200 Košice, Slovak Republic

B. M. VINAGRE**
Department of Electronics and Electromechanical Engineering, Industrial Engineering School, University of Extramadura, E-06071 Badajoz, Spain

P. O'LEARY***
Institute of Automation, Montanuniversität Leoben, A-8700 Leoben, Austria

L'. DORČÁK*
Department of Informatics and Process Control, BERG Faculty, Technical University of Košice, 04200 Košice, Slovak Republic

(Received: 16 July 2001; accepted: 7 December 2001)



Analogue realization

$$Z(s)=Z_1(s)+\cfrac{1}{Y_2(s)+\cfrac{1}{Z_3(s)+\cfrac{1}{Y_4(s)+\cfrac{1}{\cdots\cdots\cdots\cdots\cdots\cdots \atop \cfrac{1}{Y_{2n-2}(s)+\cfrac{1}{Z_{2n-1}(s)+\cfrac{1}{Y_{2n}(s)}}}}}}}$$

$$Z(s)=\cfrac{1}{Z_1(s)+\cfrac{1}{Y_2(s)+\cfrac{1}{Z_3(s)+\cfrac{1}{\cdots\cdots\cdots\cdots\cdots\cdots \atop \cfrac{1}{Y_{2n-2}(s)+\cfrac{1}{Z_{2n-1}(s)+\cfrac{1}{Y_{2n}(s)}}}}}}}$$

Fractor: Analogue device

Fractional Calculus Day at USU, April 19, 2005

Identification of a VO system

Eur. Phys. J. Special Topics **193**, 93–104 (2011)
© EDP Sciences, Springer-Verlag 2011
DOI: 10.1140/epjst/e2011-01384-4

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Regular Article

A Physical experimental study of variable-order fractional integrator and differentiator

H. Sheng[1,3], H.G. Sun[2,3], C. Coopmans[3], Y.Q. Chen[3,a], and G.W. Bohannan[3]

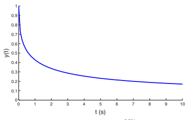


Fig. 1. Numerical result of $D_t^{-q(t)}1(t)$.

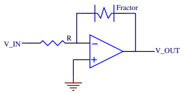


Fig. 2. Schematic circuit diagram for a fractional-order integrator.

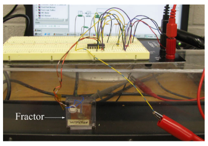


Fig. 3. Experiment realization for fractional order integrator.

Variable-order fractional differentiation and integration (VO-FD, VO-FI)

Integration and differentiation to a variable fractional order
Integral Transforms and Special Functions
Volume 1, Issue 4, 1993, Pages 277 - 300
Authors: Stefan G. Samko[a]; Bertram Ross[a]
DOI: 10.1080/10652469308819027

$$\mathfrak{D}_{c+}^{-q(t)} f(t) = \frac{1}{\Gamma[q(t)]}\int_c^t (t-\sigma)^{q(t)-1} f(\sigma)\,d\sigma.$$

$$D_{c+}^{q(t)} f(t) = \frac{1}{\Gamma[1-q(t)]}\frac{d}{dt}\int_c^t \frac{f(\sigma)}{(t-\sigma)^{q(t)}}\,d\sigma.$$

Hindawi Publishing Corporation
International Journal of Differential Equations
Volume 2010, Article ID 846107, 16 pages
doi:10.1155/2010/846107

Research Article

On the Selection and Meaning of Variable Order Operators for Dynamic Modeling

Lynnette E. S. Ramirez and Carlos F. M. Coimbra

School of Engineering, University of California, P.O. Box 2039, Merced, CA 95344, USA

Correspondence should be addressed to Carlos F. M. Coimbra, ccoimbra@ucmerced.edu

$${}^{1}D_0^{-q(t)} f(t) = \int_0^t \frac{(t-\sigma)^{q(t)-1}}{\Gamma[q(t)]} f(\sigma)\,d\sigma,$$

$${}^{2}D_0^{-q(t)} f(t) = \int_0^t \frac{(t-\sigma)^{q(\sigma)-1}}{\Gamma[q(\sigma)]} f(\sigma)\,d\sigma,$$

$${}^{3}D_0^{-q(t)} f(t) = \int_0^t \frac{(t-\sigma)^{q(t-\sigma)-1}}{\Gamma[q(t-\sigma)]} f(\sigma)\,d\sigma.$$

$${}^{3-5}D_0^{q(t)} f(t) = \frac{d}{dt}\left(\int_0^t \frac{(t-\sigma)^{q(t,\sigma)-1}}{\Gamma[q(t,\sigma)]} f(\sigma)\,d\sigma\right),$$

$${}^{6-8}D_0^{q(t)} f(t) = \int_0^t \frac{(t-\sigma)^{q(t,\sigma)-1}}{\Gamma[q(t,\sigma)]} f^{(1)}(\sigma)\,d\sigma,$$

$${}^{9}D_0^{-q(t)} f(t) = \frac{1}{\Gamma[1-q(t)]}\int_{0+}^t (t-\sigma)^{-q(t)} f^{(1)}(\sigma)\,d\sigma + \frac{(f(0+)-f(0-))t^{-q(t)}}{\Gamma[1-q(t)]},$$

Identification of a VO system

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY 1

Experimental Evidence of Variable-Order Behavior of Ladders and Nested Ladders

Dominik Sierociuk, *Member, IEEE*, Igor Podlubny, *Member, IEEE*, and Ivo Petráš, *Member, IEEE*

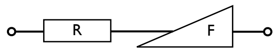
***Abstract*—The experimental study of two kinds of electrical circuits, a domino ladder and a nested ladder, is presented. While the domino ladder is known and already appeared in the theory of fractional-order systems, the nested ladder circuit is presented in this article for the first time. For fitting the measured data, a new approach is suggested, which is based on using the Mittag-Leffler function and which means that the data are fitted by a solution of an initial-value problem for a two-term fractional differential equation. The experiment showed that in the frequency domain the domino ladder behaves as a system of order 0.5 and the nested ladder as a system of order 0.25, which is in perfect agreement with the theory developed for their design. In the time domain, however, the order of the domino ladder is changing from roughly 0.5 to al-**

$[a,b]$ is ${}_aD_t^{\alpha} f(t)$, with $\alpha \in \mathbb{R}$. Sometimes a simplified notation $f^{(\alpha)}(t)$ or $d^{\alpha} f(t)/dt^{\alpha}$ is used. In some applications also right-sided fractional derivatives ${}_tD_b^{\alpha} f(t)$ are used, but in the present article we will use only left-sided fractional derivatives. Even from the notation one can see that evaluation of the left-sided fractional-order operators require the values of the function $f(t)$ in the interval $[a,t]$. When α becomes an integer number, this interval shrinks to the vicinity of the point t, and we obtain the classical integer-order derivatives as particular cases.

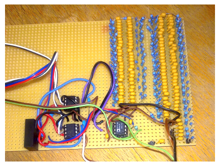
There are several definitions of the fractional derivatives and integrals, of which we need only the following two.

Identification of a VO system

The system:



F (fractance with α = 0.5): domino ladder with 60 or 130 steps

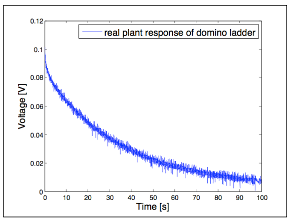


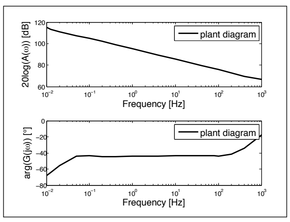
Equipment used for measurements:

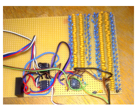
- dSPACE ACE Kit 1103
- PX4 Expansion Box
- NET KIT (Ethernet interface)
- RTI CAN Blockset
- CLP1103 (connector panel)

- Matlab
- Real Time Toolbox.

Identification of a VO system





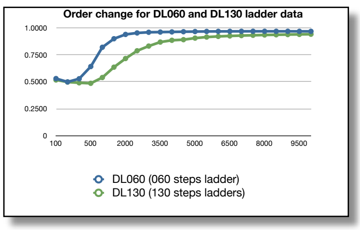


130 steps in this ladder
(also tried ladder
with 60 steps)

Identification of a VO system

Fitting by $y = y_0\, E_{\alpha,1}(a\, t^{\alpha})$

in subintervals $[0, t_n]$.



n	alpha DL060	alpha DL130
100	0.5294	0.5167
200	0.4984	0.4972
300	0.5277	0.4901
500	0.6408	0.4855
1000	0.8195	0.5390
1500	0.8986	0.6350
2000	0.9385	0.7152
2500	0.9523	0.7870
3000	0.9586	0.8306
3500	0.9604	0.8677
4000	0.9620	0.8842
4500	0.9638	0.8900
5000	0.9651	0.9044
5500	0.9661	0.9142
6000	0.9668	0.9201
6500	0.9670	0.9246
7000	0.9674	0.9279
7500	0.9678	0.9307
8000	0.9677	0.9336
8500	0.9676	0.9354
9000	0.9677	0.9373
9500	0.9672	0.9388
10000	0.9672	0.9406

Conclusions

- Fitting data using the Mittag-Leffler function is a natural step in approximation
- Fitting data using the Mittag-Leffler function is a kind of "autotuning fitting" – it has capability of uncovering those characteristics of the data which may be unnoticed by a human
- Fitting data using the M-L function can be used for identification of fractional-order and variable-order systems
- Fitting data using the M-L function can be used for re-formulation of basic laws in various fields of science

Our ML menu for take-away:

1. Collect your experimental data
2. Fit them using the Mittag-Leffler function

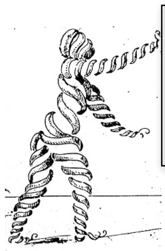
$$y = y_0 \, t^{\beta-1} E_{\alpha,\beta}(a\, t^{\alpha})$$

3. Take away your fractional-order model

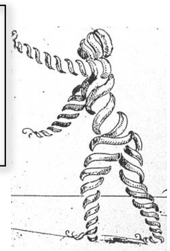
$${}_0^C D_t^{\alpha} y(t) - k\, y(t) = 0, \quad y(0) = y_0$$

But:

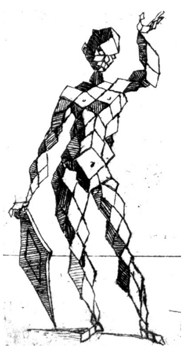
Which criterion to use for fitting?



Least Squares Revisited

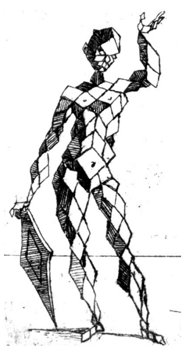


Outline



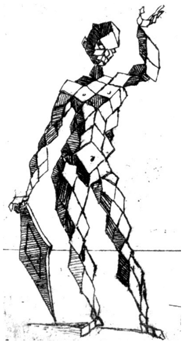
- The method
  - History of the least squares method
  - Why squares?
  - The method of least circles!
  - Arguments in favor of orthogonal distance regression
- State space description of national economies
- Results for Scandinavian and V4 countries

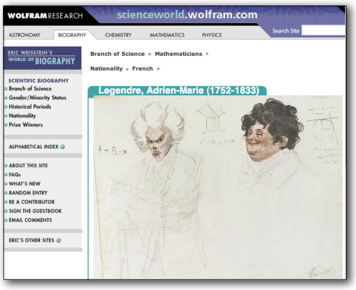
History of Least Squares



- 1805: A.-M. Legendre, "Nouvelles methodes pour la determination des cometes" — introduced and named the method,
- 1809: C. F. Gauss, "Theoria motus corporum coelestium in sectionibus conicis solem ambientium" — mentioned Legendre's work, stated that he himself was using the method since 1795.
- Legendre wrote a letter saying that claims of priority should not be made without a proof by previous publications.

BTW: meet real Legendre

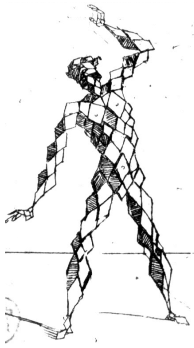




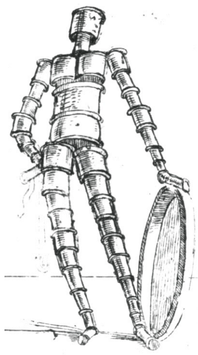
Full story: see *Notices of AMS*, Dec 2009, Vol. 56, no. 11

History of Least Squares

- Gauss did not have such a publication.
- His own computational notes were lost.
- His diary entry of 1798 on probability theory different from Laplace's is unclear.
- His colleagues apparently did not remember discussions with Gauss on this method.

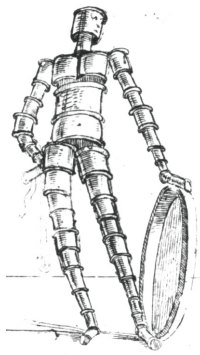


History of Least Squares



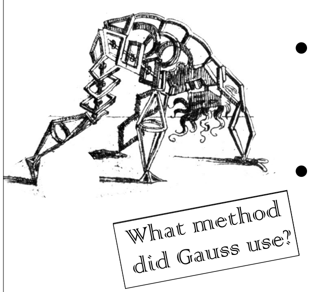
- 1816: astronomer H. Olbers included in his paper a footnote asserting that Gauss had shown him LSM in 1802.
- 1832: a similar note made by another astronomer, F.W. Bessel.

History of Least Squares



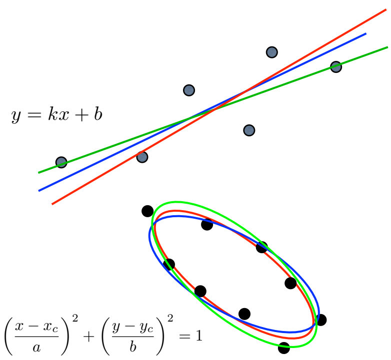
- 1831: H. Schumacher suggested repeating the calculations from Gauss's 1799 paper to demonstrate that the method of least squares was indeed used by Gauss in 1799.
- Gauss did not permit this, and wrote that his word should be enough.

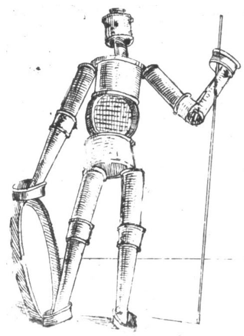
History of Least Squares



- 1981: S. M. Stigler repeated Gauss's calculations. He could not reproduce Gauss's results.
- 1998: A. Celmins reviewed the adjustments suggested by Stigler and conclude that the results published by Gauss certainly are not obtained by a minimization of observational errors in a least-squares sense.

Fitting points to lines

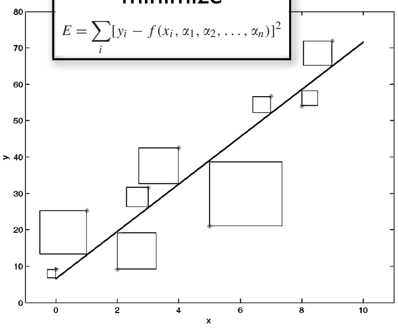




Classical Least Squares

minimize

$$E = \sum_i [y_i - f(x_i, \alpha_1, \alpha_2, \ldots, \alpha_n)]^2$$



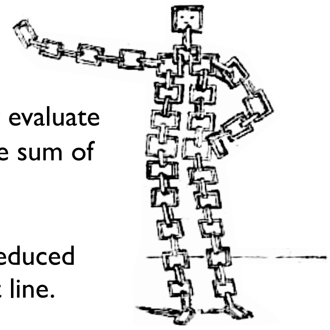
Why the Squares ?

- The simplest case is

  $y = kx + b$

- It is easy to analytically evaluate k and b to minimize the sum of squares of offsets of y.

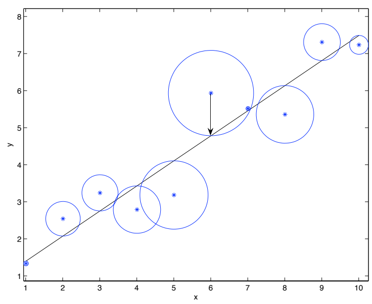
- Many other cases are reduced to the case of a straight line.

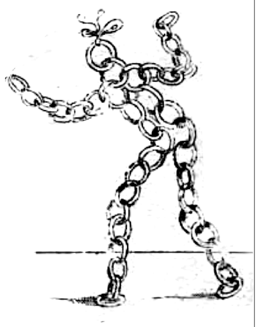


The Method of Least Circles!

$$E = \sum_i [y_i - f(x_i, \alpha_1, \alpha_2, \ldots, \alpha_n)]^2$$

$$E = \sum_i \pi [y_i - f(x_i, \alpha_1, \alpha_2, \ldots, \alpha_n)]^2$$

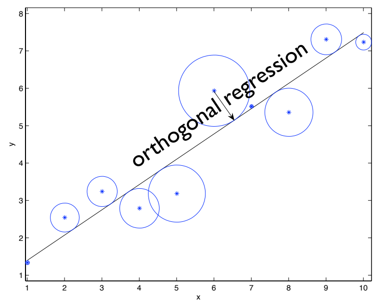


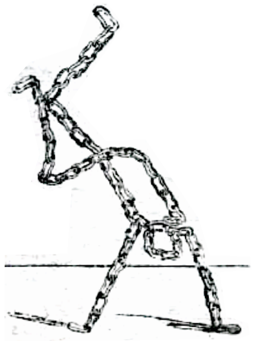


The Method of Least Circles!

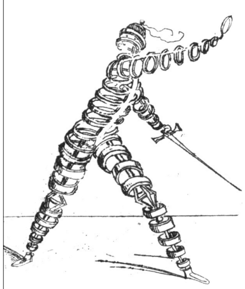
minimize squares of distances

$$E = \sum_i \pi \left[d((x_i, y_i), f(x, \alpha_1, \alpha_2, \ldots, \alpha_n)) \right]^2$$

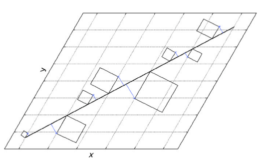




Arguments for orthogonal regression

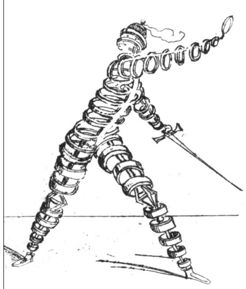


1. The shortest (orthogonal) distance is the most natural viewpoint on any fitting.
2. The sum of orthogonal distances is invariant with respect to the choice of the system of coordinates.



The squares in non-Cartesian coordinates have even less meaning.

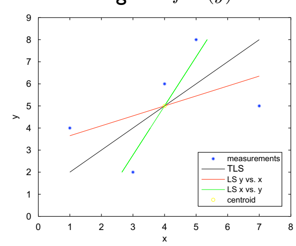
Arguments for orthogonal regression



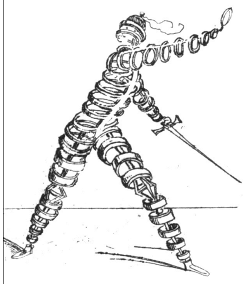
3. There are no conjugate regression lines, which appear after swapping x and y, because in the case of orthogonal regression the fitting $y = f(x)$ gives exactly the same line as the fitting $x = f^{-1}(y)$.

Nievergelt's example (1994)

x	1	3	4	5	7
y	4	2	6	8	5



Arguments for orthogonal regression



4. There are no problems with causality (normally, determination of what is an independent variable and what is a dependent variable is simply unclear or even impossible; this is always postulated).

5. Implementation of the orthogonal fitting does not depend on the number of dimensions.

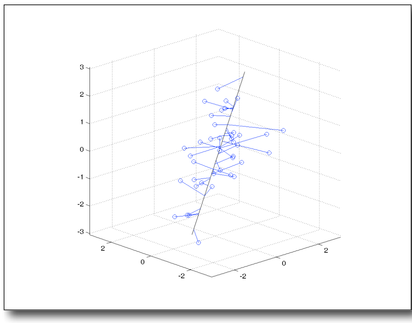
$P(x_1, x_2, \ldots, x_n)$

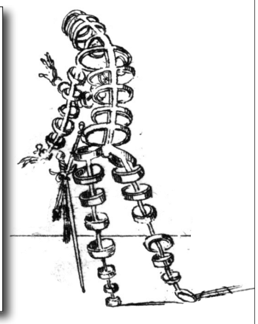
$Q(y_1, y_2, \ldots, y_n)$

$$[d(P,Q)]^2 = \sum_{i=1}^{n} (x_i - y_i)^2.$$

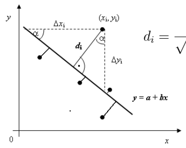
"The Flight of the Bumblebee"

Fitting 3D data by a straight line in 3D

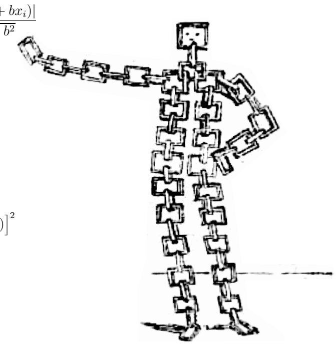




Orthogonal Linear Regression



$$d_i = \frac{\Delta y_i}{\sqrt{1+\tan^2\alpha}} = \frac{|y_i - (a + bx_i)|}{\sqrt{1+b^2}}$$



Following Legendre

$$E_\perp^2 = \sum_{i=1}^{n} \frac{[y_i - (a+bx_i)]^2}{1+b^2} = \frac{1}{1+b^2}\sum_{i=1}^{n} [y_i - (a+bx_i)]^2$$

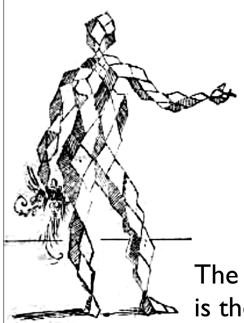
$$\frac{\partial E_\perp^2}{\partial a} = 0, \qquad \frac{\partial E_\perp^2}{\partial b} = 0,$$

we have two solutions:

$$y = a_1 + b_1 x, \qquad y = a_2 + b_2 x \qquad b_1 b_2 = -1$$

Take that b that gives the smaller criterion value.

State space for economies



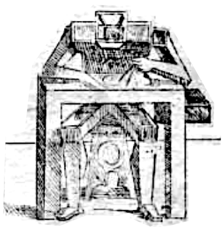
State variables $x(t), y(t), z(t)$:

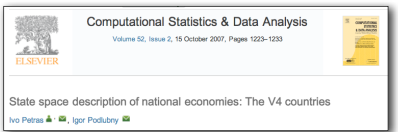
- GDP rate
- unemployment rate
- inflation rate

State of economy at $t = t_*$ is described by $\{x(t_*), y(t_*), z(t_*)\}$

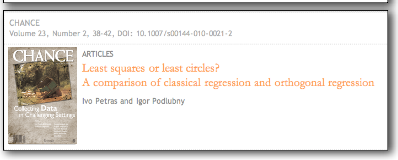
The set of points $\{x(t), y(t), z(t)\},\ t \in [T_1, T_2]$ is the trajectory of the economy.

Further reading



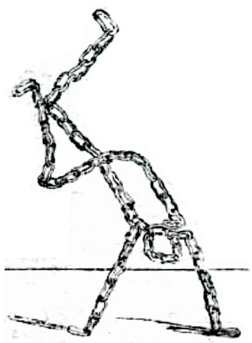

Computational Statistics & Data Analysis
Volume 52, Issue 2, 15 October 2007, Pages 1223–1233

State space description of national economies: The V4 countries
Ivo Petras, Igor Podlubny


CHANCE
Volume 23, Number 2, 38-42, DOI: 10.1007/s00144-010-0021-2

ARTICLES
Least squares or least circles?
A comparison of classical regression and orthogonal regression
Ivo Petras and Igor Podlubny

Thank you!



Illustrations:

Giovanni Battista Braccelli, Bizzarie di Varie Figure, Lisvorno, 1624.
(available at: The L. J. Rosenwald Collection, #1345, Library of Congress)