Efficient Parameter-Estimating Algorithms for Symmetry-Motivated Models: Econometrics and Beyond

Vladik Kreinovich[1], Anh H. Ly[2], Olga Kosheleva[1]
and Songsak Sriboonchitta[3]

[1]University of Texas at El Paso, USA
olgak@utep.edu, vladik@utep.edu
[2]Banking University of Ho Chi Minh City, 56 Hoang Dieu 2
Quan Thu Duc, Thu Duc, Ho Chí Minh City
[3]Faculty of Economics, Chiang Mai University
Chiang Mai 50200 Thailand, songsak econ@gmail.com

1. Need for Prediction

- In many real-life situations, we have a quantity x that changes with time t.
- We want to use the previous values of this quantity to predict its future values.
- For example:
  - we know how the stock price has changed with time, and
  - we want to use this information to predict future stock prices.
- In many cases, such a prediction is possible; for example:
  - when weather records show clear yearly cycles,
  - it is reasonable to predict that a similar yearly cycle will be observed in the future as well.

Home Page
Title Page
◀◀ ▶▶
◀ ▶
Page 2 of 39
Go Back
Full Screen
Close
Quit

2. How Can We Predict: Main Idea

- A usual approach to prediction is that we select some *model*, i.e., some parametric family of functions

$$f(t, c_1, \ldots, c_\ell).$$

- Based on the available observations, we find the parameters $\widetilde{c}_i$ which provide the best fit.
- hen we use these values $\widetilde{c}_j$ to predict the future values of the quantity x as $x(t) \approx f(t, \widetilde{c}_1, \ldots, \widetilde{c}_\ell)$.

3. Examples of Models

- In some cases, the dependence of the quantity x on time t is polynomial, in which case

$$f(t, c_1, \ldots, c_\ell) = c_1 + c_2 \cdot t + c_3 \cdot t^2 + \ldots + c_\ell \cdot t^{\ell-1}.$$

- For a simple periodic process, the dependence of the quantity x on time is described by a sinusoid:

$$f(t, c_1, c_2, c_3) = c_1 \cdot \sin(c_2 \cdot t + c_3).$$

- To get a more realistic description of a periodic process, we need to take into account higher harmonics:

$$f(t, c_1, c_2, \ldots) = c_1 \cdot \sin(c_2 \cdot t + c_3) + c_4 \cdot \sin(2c_2 \cdot t + c_5) + \ldots$$

- For a simple radioactive decay, the amount of radioactive material decreases exponentially:

$$f(t, c_1, c_2) = c_1 \cdot \exp(-c_2 \cdot t).$$

4. Examples of Models (cont-d)

- A more realistic model is a mixture of several different isotopes, with different half-lives:

  $$f(t, c_1, c_2, \ldots) = c_1 \cdot \exp(-c_2 \cdot t) + c_3 \cdot \exp(-c_4 \cdot t) + \ldots$$

- Other models include *log-periodic model* which is used to predict economic crashes:

  $$c_1 + c_2 \cdot (c_3 - t)^{c_4} + c_5 \cdot (c_3 - t)^{c_4} \cdot \cos(c_6 \cdot \ln(c_3 - t) + c_7).$$

- The following software model describes the number of bugs discovered by time t:

  $$f(t, c_1, c_2, c_3) = c_1 \cdot \ln(t - c_2) + c_3.$$

- A more complex example is a neural network, when c_j are the corresponding weights.

5. How Do We Estimate the Parameters?

- Usually, the Least Squares method is used to estimate the values of the parameters $c_1, \ldots, c_\ell$.
- So, based on the observed values $x(t_i)$, we find c_j that minimize $\sum_{i=1}^{n} (x_i - f(t_i, c_1, \ldots, c_\ell))^2$.
- In some cases – e.g., for the polynomial dependence – the model $f(x, c_1, \ldots, c_\ell)$ linearly depends on c_j.
- Then, the minimized expression is quadratic in c_j.
- We can find the minimum of a function of several variables by equating all its partial derivatives to 0.
- For a quadratic objective function, all the partial derivatives are linear functions of c_j.

6. How Do We Estimate the Parameters (cont-d)

- Thus, by equating them all to 0, we get a system of linear equations for the unknowns c_j.
- For solving systems of linear equations, there are many efficient algorithms.
- So in this case, the problem of identifying the model's parameters is computationally easy.
- On the other hand, in general, the dependence of the model on the parameters c_j is non-linear.
- Thus, the objective function is more complex than quadratic.
- It is known that, in general, optimization is computationally intensive – NP-hard.
- It is therefore desirable to select models for which identification is easier. But how do we select modles?

7. How Are Models Selected in the First Place?

- Sometimes, we have an good understanding of the processes that cause the quantity x to change.
- In such situations, we have a theoretically justified model.
- In most cases, however, the model is selected empirically:
  - we try different models, and
  - we select the one for which, for the same number of parameters, the approximation error is min.

8. Often, the Empirical Efficiency of Selected Models Can Be Explained by Symmetry

- In an empirical choice, we only compare a few possible models.
- As a result
  - the fact that the selected model turned out to be better than others
  - does not necessarily mean that this model is indeed the best for a given phenomenon:
  - there are, in principle, many other models that we did not consider in our empirical comparison.
- Good news is that in many cases, the empirical selection can be confirmed by a theoretical analysis.
- Often, the empirically successful model can be derived from the natural symmetry requirements.

9. But the Model Remains Computationally Intensive

- The fact that the empirically selected model is theoretically justified does not change its formulas; so:
  - if the dependence of this model on the corresponding parameters c_j is non-linear,
  - the problem of identifying parameters of this model remains computationally intensive.
- In this talk, we show that symmetries:
  - are not only helpful in selecting a model,
  - they can also help design computationally efficient algorithms for identifying model's parameters.

10. How Symmetries Justify Models: A Brief Reminder

- In some practical cases, the changes in the quantity x come from a single and simple process.
- This is the situation, e.g., with most oscillations.
- In most practical cases, however, many different factors lead to changes in x.
- Some of these changes are independent, and may have different intensity.
- Thus, $x(t)$ can be represented as a linear combination of the different factors:

$$C_1 \cdot e_1(t) + \ldots + C_m \cdot e_m(t) \text{ for some } e_j(t).$$

- This is the case for polynomials, when $e_1(t) = 1$, $e_2(t) = t$, $e_3(t) = t^2$, etc.

11. How Symmetries Justify Models (cont-d)

- This is the case for periodic processes, when:
  - $e_1(t)$ is the main sinusoid,
  - $e_2(t)$ is the sinusoid corresponding to double frequency,
  - $e_3(t)$ is the sinusoid corresponding to triple frequency, etc.
- This is the case for radioactive decay, where $e_j(t)$ are exponential functions with different hall-life.
- In all these cases, $e_j(t)$ are differentiable (smooth).
- So, without losing generality, we can assume that these functions are smooth.
- In these terms, selecting a model means selecting the corresponding functions $e_1(t), \ldots, e_m(t)$.

12. What Natural Symmetries Should We Consider?

- Many physical processes – such as radioactive decay – do not have a starting point.
- Their general properties do not change:
  - whether we consider the piece of a radioactive material now
  - or in a hundred years.
- The exact amount of the material will decrease.
- However, its properties – and its rate of decay – will remain the same.

13. What Natural Symmetries (cont-d)

- In such situations:
  - the observed value $x(t)$ changes with time, but
  - the whole family of functions should not change
  - if we simply start counting time from a different starting point.
- If we start to count time from a starting point which is t_0 moments in the future, then:
  - moment t in the new scale
  - corresponds to moment $t + t_0$ in the original scale.
- Thus:
  - if in the new scale, the set of functions has the above form,
  - then these same functions in the original time scale have the form $C_1 \cdot e_1(t + t_0) + \ldots + C_m \cdot e_m(t + t_0)$.

14. What Natural Symmetries (cont-d)

- The above natural requirement then says that the two families must coincide – i.e., that:
  - every function from the new family can be expressed in the old form (with different C_j),
  - and vice versa, every function from the old family can be expressed in the old form.
- In other cases,
  - there *is* a natural starting (or ending) point t_0, but
  - there is no preferred time unit.
- In such cases, it is reasonable to require that:
  - if we use a different unit for measuring time,
  - nothing will change,
  - in particular, the class of possible dependencies should not change.

15. What Natural Symmetries (cont-d)

- If we keep t_0 as the starting point, and use a measuring unit which is λ times smaller, then we get

$$t' = t_0 + \lambda \cdot (t - t_0).$$

- It is therefore reasonable to require that:
  - if we make this change,
  - the family of approximating functions remains the same.
- The new family has the form.

$$C_1 \cdot e_1(t_0 + \lambda \cdot (t - t_0)) + \ldots + C_m \cdot e_m(t_0 + \lambda \cdot (t - t_0)).$$

- The new family must coincides with the original family.

16. What Can We Conclude From These Symmetry Requirements

- We will consider the two cases separately.
- First, the case of shift-invariance.
- Then, the case of scale-invariance.

17. Case of Shift-Invariance

- In the shift-invariant case, every shifted function also belongs to the original family.
- In particular, for every j and t_0, we have:

$$e_j(t+t_0) = C_{1j}(t_0)\cdot e_1(t) + \ldots + C_{mj}(t_0)\cdot e_m(t).$$

- For each t, we can consider the equation (5) at m different moments of time $t = t_1, \ldots, t_m$.
- Then, we get the following system of m linear equations with m linear unknowns $C_{1j}(t_0), \ldots, C_{mj}(t_0)$:

$$e_j(t_1+t_0) = C_{1j}(t_0)\cdot e_1(t_1) + \ldots + C_{mj}(t_0)\cdot e_m(t_1),$$

$$e_j(t_2+t_0) = C_{1j}(t_0)\cdot e_1(t_2) + \ldots + C_{mj}(t_0)\cdot e_m(t_2),$$

$$\ldots$$

$$e_j(t_m+t_0) = C_{1j}(t_0)\cdot e_1(t_m) + \ldots + C_{mj}(t_0)\cdot e_m(t_m).$$

18. Case of Shift-Invariance (cont-d)

- The solution to a linear system can be explicitly described by the Cramer's rule.
- According to this rule, the solution is a ratio of two determinants.
- So, the solution is a differentiable function of the right-hand sides and of the coefficients at the unknowns.
- Since the functions $e_j(t)$ are smooth, the right-hand sides and the coefficients are also smooth.
- Thus, thus the solution $C_{j'j}(t_0)$ is a differentiable function of differentiable functions.
- It is, thus, a smooth function itself.
- Since $e_{j'}(t)$ and $C_{j'j}(t_0)$ are differentiable, we can differentiate the equations by t_0 and take $t_0 = 0$:

$$e'_j(t) = c_{1j} \cdot e_1 + \ldots + c_{mj} \cdot e_m, \text{ where } c_{j'j} \stackrel{\text{def}}{=} C'_{j'j}(0).$$

19. Case of Shift-Invariance (cont-d)

- Thus, $e_1(t), \ldots, e_m(t)$ satisfy a system of m linear differential equations with constant coefficients.
- A general solution to this system of equations is well known.
- It is a linear combination of functions of the type $t^k \cdot \exp(\lambda \cdot t)$, where λ are eigenvalues of the matrix $c_{j'j}$.
- Factors $t, t^2, \ldots, t^q$ appear if the corresponding eigenvalue is multiple, with multiplicity q.
- Please note that the eigenvalues are, in general, complex numbers $\lambda = a + b \cdot \mathrm{i}$, in which case

$$\exp(\lambda \cdot t) = \exp(a \cdot t) \cdot (\cos(b \cdot t) + \mathrm{i} \cdot \sin(b \cdot t)).$$

- In real-valued terms, each function $e_j(t)$ is thus a linear combination of functions of the type

$$t^k \cdot \exp(a \cdot t) \cdot (\cos(b \cdot t) + \mathrm{i} \cdot \sin(b \cdot t)).$$

20. Case of Scale-Invariance (cont-d)

- Let us now consider the case of scale-invariance with respect to the special point t_0.
- To simplify our analysis, let us consider, instead of time, an auxiliary variable $\tau \stackrel{\text{def}}{=} \ln(t - t_0)$.
- In terms of this auxiliary variable, we have $t = t_0 + \exp(\tau)$, and the original functions $e_i(t)$ take the form

$$E_i(\tau) = e_i(t_0 + \exp(\tau)).$$

- In terms of the new variable τ, the scaling transformation takes the form $\tau \to \tau + \tau_0$, where $\tau_0 \stackrel{\text{def}}{=} \ln(\lambda)$.

21. Case of Scale-Invariance (cont-d)

- Thus, scale-invariance means that:
  - the original class of functions

$$C_1 \cdot E_1(\tau) + \ldots + C_m \cdot E_m(\tau)$$

  - coincides with the transformed family

$$C_1 \cdot E_1(\tau + \tau_0) + \ldots + C_m \cdot E_m(\tau + \tau_0).$$

- So, each $E_j(\tau)$ is a linear combination of functions $\tau^k \cdot \exp(\lambda \cdot \tau) = \tau^k \cdot \exp(a \cdot \tau) \cdot (\cos(b \cdot \tau) + \mathrm{i} \cdot \sin(b \cdot \tau)).$
- We can substitute $\tau = \ln(t - t_0)$ into this formula.
- So, we conclude that each function $e_j(t)$ is a linear combination of functions of the type

$$(\ln(t-t_0))^k \cdot (t-t_0)^\lambda =$$

$$(\ln(t-t_0))^k \cdot (t-t_0)^a \cdot (\cos(b \cdot \ln(t-t_0)) + \mathrm{i} \cdot \sin(b \cdot \ln(t-t_0))).$$

22. Comments

- These formulas are highly non-linear.
- So, it is computationally difficult to identify the parameters of these models from observations.
- What if we have both shift- and scale-invariance?
- In this cases, the expression should be both:
  - a linear combination of the terms $t^k \cdot \exp(\lambda \cdot t)$ and
  - a combination of the terms of the type

$$(\ln(t - t_0))^k \cdot (t - t_0)^\lambda.$$

- The need for the second interpretation excludes exponential terms.
- So, such functions should be linear combinations of terms x^k, i.e., polynomials.

23. Comments (cont-d)

- This is the only case when the dependence on the parameters is linear and so, computationally easy.
- Let us described how to make identification of the parameters of these models easy.

24. Computationally Efficient Parameter Identification: Main Idea

- We would like to come up with a linear differential equation for symmetry-motivated models.
- To describe such an equation, let us denote the differentiation operation by D, so that $(Df)(t) \stackrel{\text{def}}{=} f'(t)$.
- Let us start with describing shift-invariant models in these terms.
- In these models, every function $e_j(t)$ is a linear combination of functions of the type $t^k \cdot \exp(\lambda \cdot t)$.
- Let us start with the case $k = 0$, when this function takes the form $\exp(\lambda \cdot t)$.
- For $\exp(\lambda \cdot t)$, we have $D \exp(\lambda \cdot t) = \lambda \cdot \exp(\lambda \cdot t)$, thus $(D - \lambda) \exp(\lambda \cdot t) = 0$.

25. Main Idea (cont-d)

- For the next ($k = 1$) function $e(t) = t \cdot \exp(\lambda t)$:

$$(De)(t) = \exp(\lambda \cdot t) + \lambda \cdot \exp(\lambda \cdot t), \text{ thus}$$

$$((D - \lambda)e)(t) = \exp(\lambda \cdot t).$$

- We already know that $(D - \lambda)\exp(\lambda \cdot t) = 0$, thus we have $((D - \lambda)^2 e)(t) = 0$.
- Similarly, for the function $e(t) = t^k \cdot \exp(\lambda \cdot t)$, we have $(De)(t) = k \cdot t^{k-1} \cdot \exp(\lambda \cdot t) + \lambda \cdot t^k \cdot \exp(\lambda \cdot t)$, thus

$$((D - \lambda)e)(t) = k \cdot t^{k-1} \cdot \exp(\lambda \cdot t).$$

- So, by induction, we can prove that for this function $e(t)$, we have $(D - \lambda)^k e = 0$.
- Different expressions forming $e_j(t)$ correspond to different eigenvalues λ_ℓ.

26. Main Idea (cont-d)

- So each of them is annihilated:
  - by a corresponding differential operation $D - \lambda_\ell$,
  - or, if this eigenvalue if multiple with multiplicity q_ℓ, by an operator $(D - \lambda_\ell)^{q_\ell}$.
- Thus, if we apply all these operators one after another, all the terms in $e_j(t)$ will be annihilated: $\widetilde{D} e_j = 0$ for

$$\widetilde{D} \stackrel{\text{def}}{=} (D - \lambda_1)^{q_1} (D - \lambda_2)^{q_2} \ldots (D - \lambda_m)^{q_m}.$$

- Since each model $x(t)$ is a linear combination of the functions $e_j(t)$, we have $\widetilde{D} x = 0$.
- If we open the parentheses, we conclude that $\widetilde{D}$ is a polynomial of m-th order in terms of D:

$$\widetilde{D} = D^m + a_1 \cdot D^{m-1} + a_2 \cdot D^{m-2} + \ldots + a_m.$$

27. Main Idea (cont-d)

- Thus, the equation $(\widetilde{D}x)(t) = 0$ takes the form

$$\frac{d^m x}{dt^m} + a_1 \cdot \frac{d^{m-1} x}{dt^{m-1}} + a_2 \cdot \frac{d^{m-2} x}{dt^{m-2}} + \ldots + a_m \cdot x = 0.$$

- This is the desired differential equation with constant coefficients.

28. Examples

- For a polynomial of order $\le m-1$, all eigenvalues are zeros, so $\widetilde{D} = D^m$.
- The corresponding differential equation is $\dfrac{d^m x}{dt^m} = 0$.
- One can see that solutions to this differential equation are indeed exactly polynomials of order $\le m-1$.
- For a simple sinusoidal signal $x(t) = A \cdot \cos(\omega \cdot t + \varphi)$, we get a second order differential equation

$$\frac{d^2 x}{dt^2} + a_1 \cdot \frac{dx}{dt} + a_2 \cdot x = 0.$$

- To be more precise, the sinusoid correspond to the case when $a_1 = 0$ and $a_2 > 0$.
- Other cases correspond to exponential functions or functions $A \cdot \exp(-a \cdot t) \cdot \cos(\omega \cdot t + \varphi)$.

29. How Can We Easily Identify a Model: Towards an Algorithm

- In terms of the original the parameters of the model, the dependence is non-linear.
- Instead, let us identify the parameters $a_1, \ldots, a_m$ of the corresponding differential equation.
- Of course, we have to approximate each derivative by a finite difference $(\Delta x)_i \stackrel{\text{def}}{=} \dfrac{x_i - x_{i-1}}{\Delta t}$.
- Then, instead of the second derivatives, we will use

$$(\Delta^2 x)_i \stackrel{\text{def}}{=} (\Delta(\Delta x))_i = \frac{(\Delta x)_i - (\Delta x)_{i-1}}{\Delta t} = \frac{x_i - 2x_{i-1} + x_{i-2}}{(\Delta t)^2}.$$

- Similarly, in the general case, we have

$$(\Delta^k x)_i = (\Delta(\Delta^{k-1} x))_i =$$
$$\frac{x_i - k \cdot x_{t-1} + C_2^k \cdot t_{i-1} - C_3^k \cdot t_{i-2} + \ldots + (-1)^k \cdot t_{i-k}}{(\Delta t)^k}.$$

30. Towards an Algorithm (cont-d)

- So, instead of the differential equation, we have an approximate equation

$$(\Delta^m x)_i + a_1 \cdot (\Delta^{m-1} x)_i + a_2 \cdot (\Delta^{m-2} x)_i + \ldots + x_i = 0.$$

- The values $(\Delta^k x)_i$ are computed based on the observations x_i.
- So, we get a system of linear equations from which we can find $a_1, \ldots, a_m$ by using the Least Squares.

31. Shift-Invariant Case: Resulting Algorithm

- Based on the sequence of observations $x_i = x(t_i)$, we compute the sequence of values $(\Delta x)_i = \dfrac{x_i - x_{i-1}}{\Delta t}$.
- Then, we compute the sequence $(\Delta^2 x)_i = (\Delta(\Delta x))_i$, etc., until we have computed $(\Delta^m x)_i$.
- We find the parameters a_j by applying the Least Squares Method to

$$(\Delta^m x)_i + a_1 \cdot (\Delta^{m-1} x)_i + a_2 \cdot (\Delta^{m-2} x)_i + \ldots + x_i = 0.$$

32. Comments

- No problem if observations are not equally spaced in time: take $(\Delta x)_i = \dfrac{x_i - x_{i-1}}{\Delta t_i}$, where $\Delta t_i \stackrel{\text{def}}{=} t_i - t_{i-1}$.
- Usually, the values $x_i = x(t_i)$ at different moments of time are uncorrelated.
- However, their linear combinations $(\Delta^j x)_i$ are correlated.
- Indeed, the expressions for i and for $i-1$ now depend on the same value x_i.
- Thus, we need to use the Least Squares in the presence of this easy-to-compute correlation.
- This does not affect the computational easiness:
  - the expression is still quadratic and
  - equating its derivatives to 0 still leads to a system of linear equations.

33. Comments (cont-d)

- If needed, we can convert the new parameters $a_1, \ldots, a_m$ into the more traditional ones.
- All we need for this is:
  - to compute the derivatives of the original expressions $f(t, c_1, \ldots, c_\ell)$ and
  - find the values a_j for which the linear combinations of these derivatives are 0s.
- Then, we get expressions describing a_j in terms of c_j:

$$a_j = f_j(c_1, \ldots, c_\ell).$$

- Once we know a_j, we can solve the corresponding system of equations $f_j(c_1, \ldots, c_\ell) = a_j$.
- This system is non-linear, but when the number of parameters is small, it is not that difficult to solve.

34. Scale-Invariant Case: Analysis of the Problem

- The scale-invariant case reduces to the shift-invariant case if we introduce an auxiliary variable $\tau = \ln(t - t_0)$.
- Thus, with respect to this new variable τ, we get a differential equation:

$$\frac{d^m x}{d\tau^m} + a_1 \cdot \frac{d^{m-1} x}{d\tau^{m-1}} + \ldots + a_m \cdot x = 0.$$

- Differentiating the relation between τ and t, we conclude that $d\tau = \dfrac{dt}{t - t_0}$.
- Thus, $\dfrac{d}{d\tau} = (t - t_0) \cdot \dfrac{d}{dt}$, and the above equation takes the form:

$$(t - t_0)^m \cdot \frac{d^m x}{dt^m} + a_1 \cdot (t - t_0)^{m-1} \cdot \frac{d^{m-1} x}{dt^{m-1}} + \ldots + a_m \cdot x = 0.$$

35. Scale-Invariant Case (cont-d)

- There are two possibilities:
  - it may be that we know t_0, or
  - it may be that we need to determine t_0 from observations.
- In the first subcase, all we need is to find the values a_j.
- In the second subcase, to make the problem linear, we expand all the polynomials

$$(t-t_0)^j = x^j + (-j\cdot t_0)\cdot t^{j-1} + \ldots$$

- Then each term $a_j\cdot(t-t_0)^{m-j}\cdot\dfrac{d^{m-j}x}{dt^{m-j}}$ becomes a linear combination of the following terms:

$$t^{m-j}\cdot\frac{d^{m-j}x}{dt^{m-j}},\quad t^{m-j-1}\cdot\frac{d^{m-j}x}{dt^{m-j}},\quad \ldots,\quad \frac{d^{m-j}x}{dt^{m-j}}.$$

36. Scale-Invariant Case (cont-d)

- Let us denote the coefficients at $t^{m-j-k}\cdot\dfrac{dx^{m-j}}{dt^{m-j}}$ by a_{jk}.
- Then, the above formula takes the following form:

$$t^m\cdot\frac{dx^m}{dt^m}+a_{01}\cdot t^{m-1}\cdot\frac{dx^m}{dt^m}+\ldots+a_{0m}\cdot\frac{dx^m}{dt^m}+$$

$$a_{10}\cdot t^{m-1}\cdot\frac{dx^{m-1}}{dt^{m-1}}+a_{11}\cdot t^{m-2}\cdot\frac{dx^{m-1}}{dt^{m-1}}+\ldots+a_{1,m-1}\cdot\frac{dx^{m-1}}{dt^{m-1}}+$$

$$\ldots+$$

$$a_{m0}\cdot x=0.$$

- Thus, depending on whether we know t_0 or we don't, we arrive at the following linear algorithms.

37. Scale-Invariant Case: Resulting Algorithms

- Based on the observations $x_i = x(t_i)$, we compute the finite differences $(\Delta^k x)_i$ for all $k \le m$.
- If we know t_0, we compute $a_1, \ldots, a_m$ of the corresponding model by applying the Least Squares to:

$$(t_i - t_0)^m \cdot (\Delta^m x)_i + a_1 \cdot (t_i - t_0)^{m-1} \cdot (\Delta^{m-1} x)_i + \ldots + a_m \cdot x_i = 0.$$

- When we do not know t_0, then we find $a_{jk}$by applying the Least Squares to:

$$t_i^m \cdot (\Delta^m x)_i + a_{01} \cdot t_i^{m-1} \cdot (\Delta^m x)_i + \ldots + a_{0m} \cdot (\Delta^m x)_i +$$

$$a_{10} \cdot t_i^{m-1} \cdot (\Delta^{m-1} x)_i + a_{11} \cdot t^{m-2} \cdot (\Delta^{m-1} x)_i + \ldots + a_{1,m-1} \cdot (\Delta^{m-1} x)_i +$$

$$\ldots$$

$$a_{m0} \cdot x = 0.$$

38. Acknowledgments

- We acknowledge the partial support of:
  - the Center of Excellence in Econometrics, Faculty of Economics,
  - Chiang Mai University, Thailand.
- This work was also supported by the Nat'l Science Found. grant HRD-1242122 (Cyber-ShARE Center).