Least-square Example - Simple Harmonic Analysis

Jan Skaloud

Sensor Orientation

There are many periodical (P) phenomena in nature or electronic instruments that can be expressed as:

$z = f(y) = f(y + P) = f(y + 2P) = ...$

In such a situation it is sufficient to obtain observation over one period, which duration $y_n - y_0 = P$ we normalize by a substitution

$t = \frac{2\pi}{P}\, y.$

to the interval $(0, 2\pi)$

Model

Let's assume that there is only one cause (e.g. an evolution of a daily or yearly temperature) that can be modeled by a sinusoidal function

$z = A_0 + a\, sin(t + A)$

where A_0 is the mean shift on the z-axis, a is the amplitude, t is the argument over the interval $(0, 2\pi)$ and $(2\pi - A)$ is the shift of the sinusoidal with respect to t origin. The initial phase for $t = 0$ corresponds to $z_0 = A_0 + a\, sinA$.

Problem

We aim to determine the unknown parameters from observations by the method of least-square, namely:

- unknowns: $x' = (A_0,\, a,\, A)$
- observations: (t_i, z_i), where $i = 1, ..., n$ with $n >= 3$

We start to form the observation equations with the help of trigonometric identity $sin(A + B) = sinA\, cosB + cosA\, sinB$

$z_i = A_0 + a(sin\, t_i\, cosA + cos\, t_i\, sinA) + v_i$

and two substitutions $a\, sinA = A_1$ and $a\, cosA = A_2$ to obtain

$z_i = A_0 + cos\, t_i \cdot A_1 + sin\, t_i \cdot A_2 + v_i$

where v_i is the correction due to random error.

Normal equations

After substitutions there is only a linear relation between observations z_i and the 3 modified unknown parameters $x = (A_0,\, A_1,\, A_2)$. We rewrite the observation equations in the vector-matrix form $z = H\, x$ where the design matrix H is obtained by stacking n of row-vectors $(1\; cos\, t_i\;\; sin\, t_i)$.

Considering all observations equally weighted (i.e. having the same precision), the solution follows from the least-square conditions $v^T v$ that forms the normal equations:

$(H^T H)\, x = H^T z$

Simplification

Let's rewrite the previous normal equations by its individual elements through inspecting the $(H^T H)$ matrix and the $(H^T z)$ vector while putting all the terms on left side:

$nA_0 + (\sum cos\, t)A_1 + (\sum sin\, t)A_2 - \sum z = 0$

$(\sum cos\, t)A_0 + (\sum cos^2\, t)A_1 + (\sum sin\, t\, cos\, t)A_2 - \sum(cos\, t \cdot z) = 0$

$(\sum sin\, t)A_0 + (\sum sin\, t\, cos\, t)A_1 + (\sum sin^2\, t)A_2 - \sum(sin\, t \cdot z) = 0$

We note that sin and cos are orthogonal functions with zero mean value per period. Hence, when the sampling interval is regular over the whole period the following sums are zero:

$\sum sin\, t = 0 \quad \sum cos\, t = 0 \quad \sum sin\, t\, cos\, t = 0$

Due to the following identities,

$\sum(cos\, at \cdot cos\, bt) = \frac{1}{2}\sum cos(a+b)\, t + \frac{1}{2}\sum cos(a-b)\, t$

$\sum(sin\, at \cdot sin\, bt) = \frac{1}{2}\sum cos(a-b)\, t - \frac{1}{2}\sum cos(a+b)\, t$

it further holds when $a = b$ ($\rightarrow cos(0) = 1$ and $\sum cos(2a\, t) = 0$) that

$\sum sin^2\, t = \sum cos^2\, t = \frac{n}{2}$,

Solution

With the previous reasoning, only the quadratic and absolute terms in the previously defined normal equations have non-zero elements. That facts leaves only one unknown per each equation:

$nA_0 - \sum z = 0 \rightarrow A_0 = \frac{1}{n}\sum z$

$\frac{n}{2}A_1 - \sum \cos\, t \cdot z = 0 \rightarrow A_1 = \frac{2}{n}\sum cos\, t \cdot z$

$\frac{n}{2}A_2 - \sum \sin\, t \cdot z = 0 \rightarrow A_2 = \frac{2}{n}\sum sin\, t \cdot z$

Parameter confidence

The empirical variances of unknowns are found by scaling the values on the main diagonal of $P = (H^T H)^{-1}$ by σ_0

$\sigma^2_{A_0} = P_{11} \cdot \sigma_0 = \frac{1}{n} \cdot \sigma_0$,

$\sigma^2_{A_1} = P_{22} \cdot \sigma_0 = \frac{2}{n} \cdot \sigma_0$,

$\sigma^2_{A_2} = P_{22} \cdot \sigma_0 = \frac{2}{n} \cdot \sigma_0$.

The empirical value of σ_0^2 is obtained from the quadratic sum of all corrections v_i divided by the degrees of freedom, which is the number of observations minus the number of unknowns

$\sigma_0 = \frac{v^T v}{n-3}$.

Remark

The original parameters (a, A) are recovered by back-substituting (A_1, A_2) into two equations $A_1 = a \sin A$, and $A_2 = a \cos A$ and solving this small system.