Linear least squares problems and smoothing filters

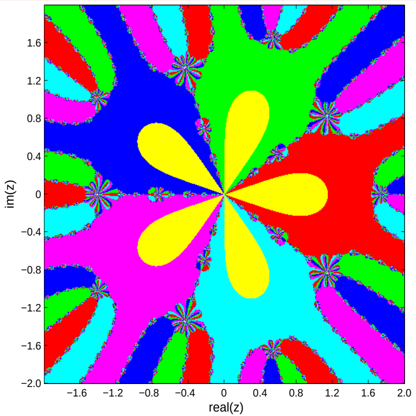
Krystyna Ziętak (Wroclaw)

Iwona Wróbel (Warsaw)

Birmingham, September 14, 2010

Outline

Goal of talk is to show that false formula can lead to better method.



Task

Given points $t_1 < \ldots < t_m$ and measured data corrupted by random noise

$$y_1, \ldots, y_m.$$

Compute smoothed value of y_i

$$\widetilde{y}_i = \sum_{j=-N_L}^{N_R} g_j y_{i+j}$$

where g_j are filter coefficients

Concept of Savitzky and Golay (1964)

interior point t_i: $1 + N_L \leq i \leq m - N_R$

polynomial $v(t)$ approximates data
over the set of points $\{t_{i-N_L}, \ldots, t_{i+N_R}\}$

in the least-squares sense

smoothed $\widetilde{y}_i = v(t_i)$

$N_L = N_R = N$

$$\widetilde{y}_i = v(t_i) = \sum_{j=-N}^{N} g_j y_{i+j} = g^T y$$

$$g = A_i (A_i^T A_i)^{-1} A_i^T e$$

A_i rectangular Vandermonde, $t_{i-N}, \ldots, t_{i+N}$

$$y = [y_{i-N}, \ldots, y_{i+N}]^T$$

$$e = [0, \ldots, 0, 1, 0, \ldots, 0]^T$$

Uniformly spaced data $t_{j+1} - t_j = \Delta$

new variable, $t \rightarrow h = (t - t_i)/\Delta$

new points $t_j \rightarrow h_j = j - i \quad t_i \rightarrow 0$

Savitzky-Golay filter coefficients

$$g^T = [g_{-N}, \dots g_N]$$

$$g = A_i(A_i^T A_i)^{-1} A_i^T e =$$
$$= B(B^T B)^{-1} B^T e = B(B^T B)^{-1} e_1$$

B rectangular Vandermonde, $-N, \dots, N$

$$e_1 = [1, 0, \dots, 0]^T, \quad e = [0, \dots, 0, 1, 0, \dots, 0]^T$$

$$B = \begin{bmatrix} 1 & -N & (-N)^2 & \cdots & (-N)^n \\ 1 & -N+1 & (-N+1)^2 & \cdots & (-N+1)^n \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & N & N^2 & \cdots & N^n \end{bmatrix}$$

$$C = \begin{bmatrix} (-N)^n & (-N)^{n-1} & \cdots & -N & 1 \\ (-N+1)^n & (-N+1)^{n-1} & \cdots & -N+1 & 1 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ N^n & N^{n-1} & \cdots & N & 1 \end{bmatrix}$$

How to compute $g = B(B^T B)^{-1} e_1$?

n degree of approx. polynomial

$$C = QR$$

$$g = \frac{1}{r_{n+1,n+1}} Q e_{n+1}$$

Gander, von Matt

Orthogonal (discrete) Gram polynomials

orthogonal over point set $\mathcal{S}_N = \{-N, \ldots, N\}$

$$q_j(h) = \sum_{k=0}^{j} (-1)^{k+j} \frac{(j+k)^{(2k)}(N+h)^{(k)}}{(k!)^2 (2N)^{(k)}}, \quad q_0(h) = 1,$$

$$a^{(0)} = 1$$
$$a^{(k)} = a(a-1)\cdots(a-k+1)$$

"discrete" Dirac delta

$$\delta_N = e = [0, \ldots, 0, 1, 0, \ldots, 0]^T \in \mathbb{R}^{n+1}$$

$$\mathcal{S}_N = \{-N, \ldots, N\}, \qquad n \quad \text{degree, even}$$

polynomial least-squares approximation over $\mathcal{S}_N$ to δ_N

$$q(h) = \frac{q_n(0) q_{n+1}(h)}{\xi \eta h}$$

ξ and η constant, depend on n and N

filter coefficients $g^T = [q(-N), \ldots, q(N)]$

orthogonal Legendre polynomials $[-1, 1]$

$$\langle P_i, P_j \rangle = \int_{-1}^{1} P_i(s) P_j(s) ds.$$

recurrence relation

$$P_k(s) = \frac{2k-1}{k} s P_{k-1}(s) - \frac{k-1}{k} P_{k-2}(s), \quad k = 1, 2, \ldots$$

$$P_0(s) = 1.$$

linear change of variable

$$[-1, 1] \to [-N, N]$$

Filter coefficients

Dirac delta $\delta(s)$

$$\delta(s) = 0 \quad \text{for } s \neq 0,$$
$$\delta(s) = \infty \quad \text{for } s = 0,$$

Persson-Strang

least-squares approximation to Dirac delta on interval $[-N, N]$ expressed by Legendre polynomials

Savitzky-Golay

least-squares approximation to vector
$e = [0, \ldots, 0, 1, 0, \ldots, 0]^T$
over the point set $\{-N, -N+1, \ldots, N\}$

Persson and Strang filter (2003)

degree of polynomial n even, $[-N, N]$

uniformly spaced data, filter coefficients $g_{-N}, \ldots, g_N$

Persson-Strang polynomial $g_j = L(j)$

$$L(h) = \frac{n+1}{2} P_n(0) \frac{P_{n+1}\left(\frac{2h}{2N+1}\right)}{h}$$

Optimal "Legendre" polynomial $g_j = K(j)$

$$K(h) = \frac{n+1}{2} P_n(0) \frac{P_{n+1}\left(\frac{h}{N}\right)}{h}$$

Examples of approximating polynomials $n = 2$

- Savitzky-Golay (Gram)

$$q(h) = \frac{-15h^2}{(2N-1)(2N+1)(2N+3)} + \frac{9N^2+9N-3}{(2N-1)(2N+1)(2N+3)}$$

- Persson-Strang

$$L(h) = \frac{-15h^2}{(2N+1)^3} + \frac{9}{4(2N+1)}$$

- optimal Legendre

$$K(h) = \frac{-15h^2}{8N^3} + \frac{9}{8N}$$

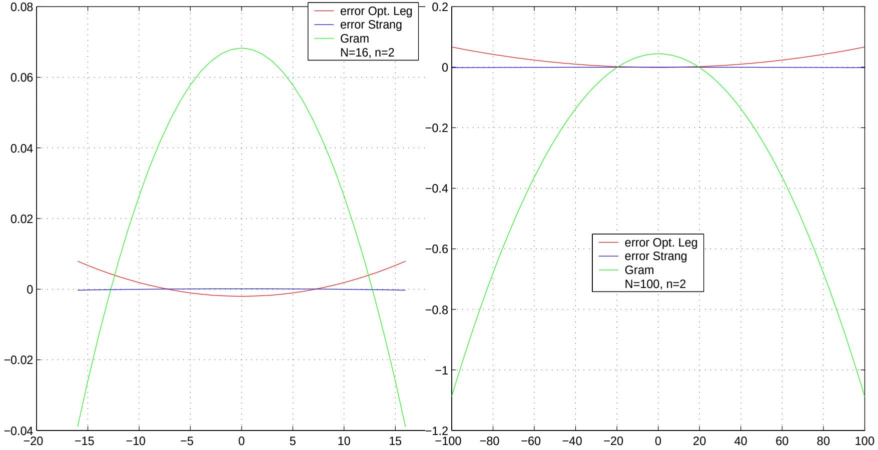


Figure : Savitzky-Golay polynomial $q(h)$ (green); differences $q(h) - L(h)$ and $q(h) - K(h)$; $N = 16$ (on left), $N = 100$ (on right); degree $n = 2$

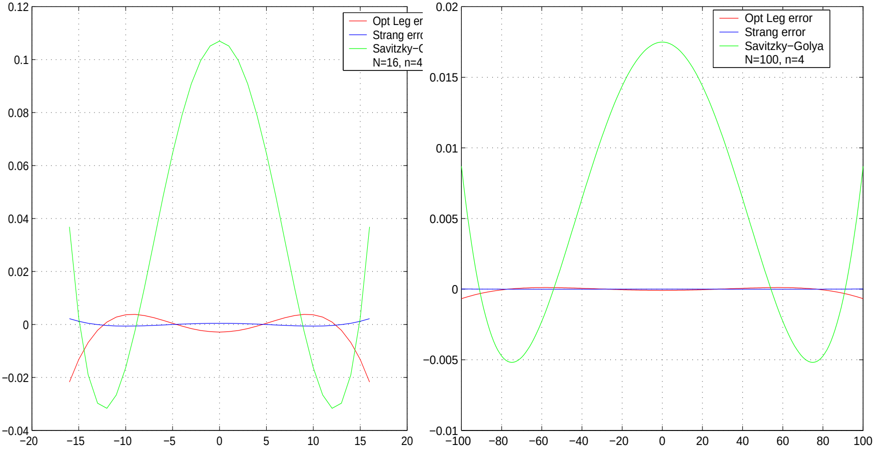


Figure : Savitzky-Golay polynomial $q(h)$ (green); differences $q(h) - L(h)$ and $q(h) - K(h)$; $N = 16$ (on left), $N = 100$ (on right); degree $n = 4$

Gander and von Matt test function on $[0, 1]$

$$F(t) = e^{-100(t-\frac{1}{5})^2} + e^{-500(t-\frac{2}{5})^2} + e^{-2500(t-\frac{3}{5})^2} + e^{-12500(t-\frac{4}{5})^2}$$

- uniform spaced data;
  $t_j = \frac{j-1}{m-1}$ for $j = 1, \ldots, m$, where $m = 1000$
- $y_j = F(t_j) + \varepsilon \times randn$, where $\varepsilon = 0.1$
- degree of polynomial $n = 4$
- window $N = 16$

Savitzky-Golay, Persson-Strang, optimal Legendre

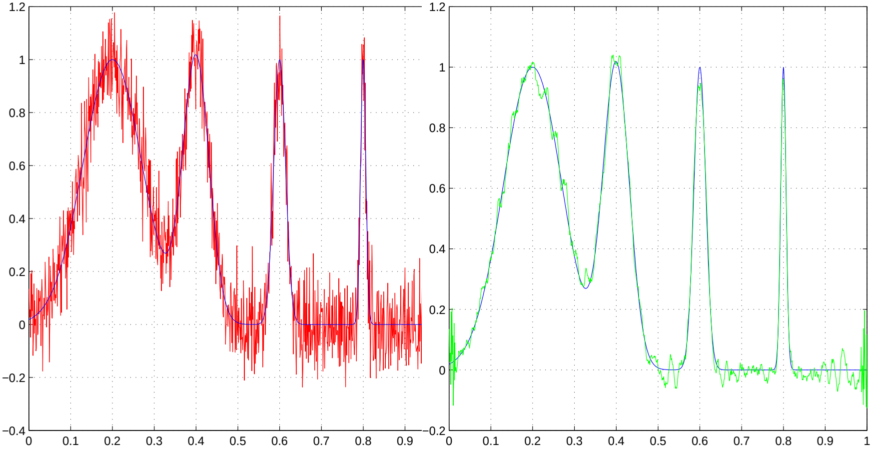


Figure : $F(t)$ exact and noise (on left); $F(t)$ exact and smoothed by Savitzky-Golay (on right)

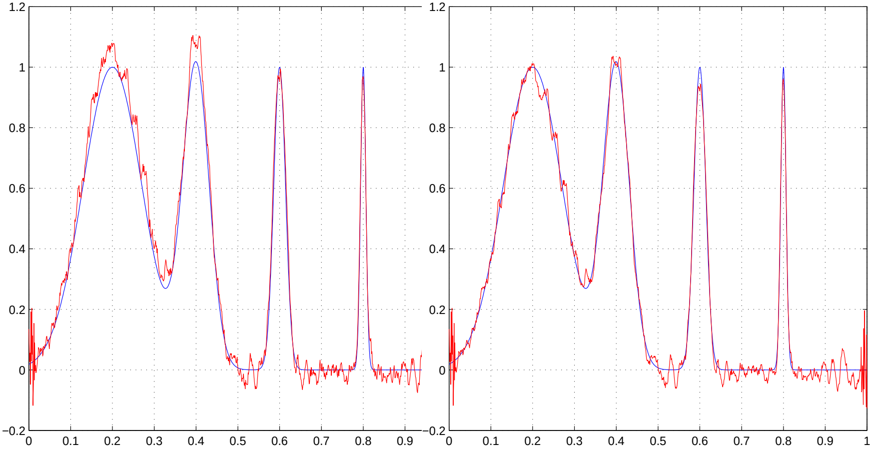


Figure : $F(t)$ exact and smoothed by optimal Legender (on left); $F(t)$ exact and smoothed by Persson-Strang (on right)

Errors: $F(t)$ exact – smoothed data (subtraction)

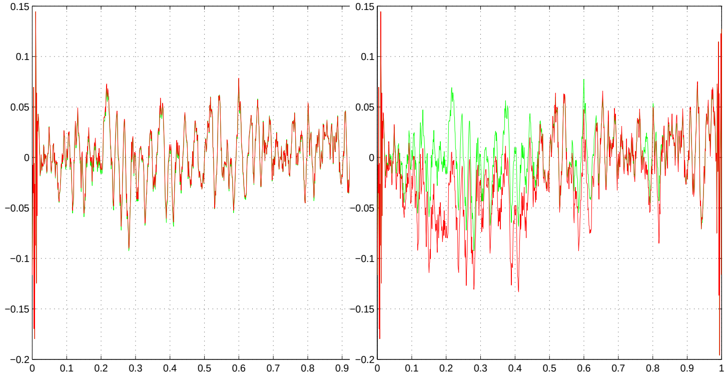


Figure : Savitzky-Golay (green) and Persson-Strang (red), (on left); Savitzky-Golay (green), optimal Legendre (red), (on right)

Gander-von Matt test function on $[0, 0.5]$

$$F(t) = e^{-100(t-\frac{1}{5})^2} + e^{-500(t-\frac{2}{5})^2} + e^{-2500(t-\frac{3}{5})^2} + e^{-12500(t-\frac{4}{5})^2}$$

- uniform spaced data;
  $t_j = \frac{j-1}{m-1}$ for $j = 1, \ldots, m$, where $m = 1000$
- $y_j = F(t_j) + \varepsilon \times randn$, where $\varepsilon = 0.1$
- degree of polynomial $n = 4$
- window $N = 32$

Savitzky-Golay, Persson-Strang, optimal Legendre

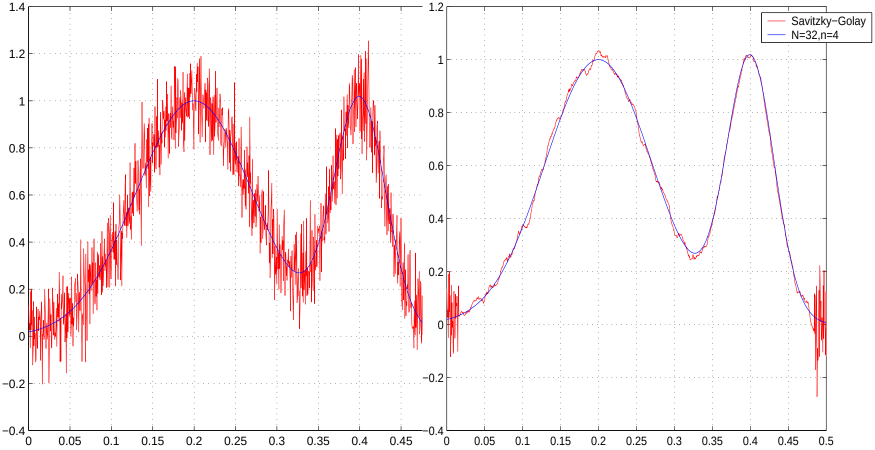


Figure : $F(t)$ exact and noise (on left); $F(t)$ exact and smoothed by Savitzky-Golay (on right); $N = 32, n = 4$

Errors: $F(t)$ exact – smoothed data (subtraction)

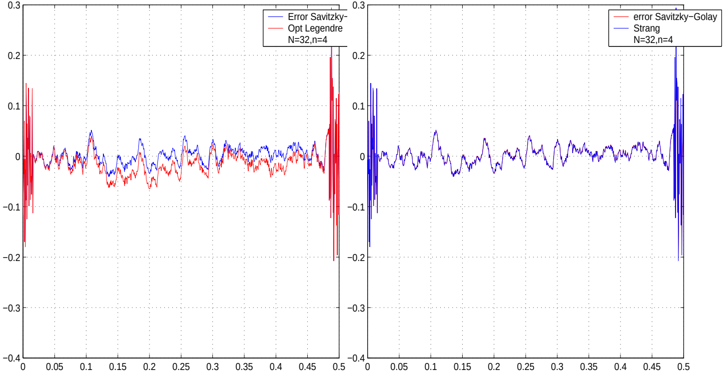


Figure : error Savitzky-Golay (blue), optim. Legendre (red), on left; error Savitzky-Golay (blue), Persson-Strang (red), on right, $N = 32, n = 4$

Persson and Strang write that the Legendre-based filters have extra advantages:

> “*in the case of irregularly spaced or missing data the polynomials stay the same and it is only the sampling points that change. (...)*
>
> The simplicity becomes especially valuable when the input no longer consists of uniformly spaced samples.
> The output from the new filter will be the natural non-uniform generalization of an ordinary convolution".

$$[t_{i-N}, t_{i+N}] \to [-N, N]$$

window $t_{i-N_L}, \ldots, t_{i+N_R}$

new variable

$$h = \frac{t - t_i}{\Delta t}, \qquad \Delta t = \frac{t_{i+N_R} - t_{i-N_L}}{N_L + N_R}$$

new points

$$h_j^{(i)} = \frac{t_j - t_i}{\Delta t} \quad \text{for } j = i - N_L, \ldots i + N_R.$$

$$\alpha_i = h_{i-N_L}^{(i)}, \qquad \beta_i = h_{i+N_R}^{(i)}$$

approximation of Dirac delta on $[\alpha_i, \beta_i]$

$P_k(t)$ Legendre polynomial

$[-1, 1] \to [\alpha_i, \beta_i]$

$$Z_k^{(i)}(h) = P_k\left(\frac{2}{\beta_i - \alpha_i}h + \frac{\alpha_i + \beta_i}{\alpha_i - \beta_i}\right)$$

$$K^{(i)}(h) = \frac{n+1}{2} \cdot \frac{Z_{n+1}^{(i)}(0)Z_n^{(i)}(h) - Z_n^{(i)}(0)Z_{n+1}^{(i)}(h)}{-h}$$

$$\tilde{y}_i = \sum_{j=i-N_L}^{i+N_R} g_j^{(i)} y_j, \quad \text{where } g_j^{(i)} = K^{(i)}(h_j^{(i)})$$

Bad news

A. Eisinberg, P. Pugliese, N. Salerno,
Numer. Math. 2001

Vandermonde matrices on integer nodes: the rectangular case

their case

$$1, 2, \ldots, N$$

- $B = V^T V$ Hankel matrix
- explicit Cholesky factor of B
- pseudo-inverse of V
- combinatorial identities

our case

$$-N, \ldots, N$$

References

Abraham Savitzky (1919–1999)
was an Americal analytical chemist. He specialized in the digital processing of infrared spectra.

Marcel J.E. Golay (1902–1989)
was a Swiss-born mathematician, physicst and information theorist.

Golay error-correcting codes

Savitzky coauthored with Marcel J.E. Golay an oft-cited paper describing the Savitzky-Golay smoothing filter.

THANK YOU FOR YOUR ATTENTION