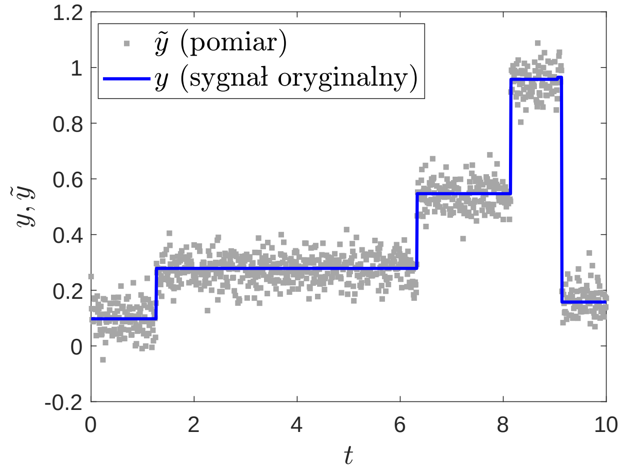
1DE2116:A - Matematyka - Metody optymalizacji EP
Rekonstrukcja sygnału kawałkami stałego. LASSO.

Maciej Twardy

Materiał przygotowany w ramach projektu „NERW 2 PW. Nauka – Edukacja – Rozwój - Współpraca". Projekt współfinansowany przez Unie Europejska w ramach Europejskiego Funduszu Społeczego. Program Operacyjny Wiedza Edukacja Rozwój 2014-2020, Os priorytetowa III Szkolnictwo Wyzsze dla gospodarki i rozwoju, Działanie 3.5 Kompleksowe programy szkół wyższych.

1 Wstęp

Będziemy rozpatrywać zagadnienie dopasowanie kawałkami stałego (ang. *piece-wise constant fitting*) [1]). Załóżmy, że dysponujemy wynikami pomiaru kawałkami stałego sygnału y. Pomiary te są obarczone szumem, zatem to co mamy do dyspozycji to wektor próbek $\tilde{y}$. Oznaczmy przez $y \in \mathbb{R}^n$ wektor próbek sygnału, przez $\tilde{y} \in \mathbb{R}^n$ wektor próbek pomiaru.



Rysunek 1: Sygnał kawałkami stały y i jego zaszumiony pomiar $\tilde{y}$.

Chcemy wyznaczyć estymatę (wektor) $\hat{y}$ sygnału y. Poszukiwanie takiej estymaty możemy sformułować jako zadanie optymalizacji (tzn. $\hat{y}$ jest rozwiązaniem rozpatrywanego zadania optymalizacji)

$$\begin{aligned} &\underset{v}{\text{minimize}} && \|\tilde{y} - v\|_2^2 && (1a) \\ &\text{subject to} && \mathbf{card}(Dv) \leqslant k && (1b) \end{aligned}$$

gdzie $\mathbf{card}(u)$ oznacza liczbę elementów niezerowych wektora u (w rozpatrywanym przypadku $u = Dv$), k jest przyjętą (arbitralnie) maksymalną liczbą zmian wartości sygnału, zaś D i Dv, odpowiednio,

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}, \quad Dv = \begin{bmatrix} v_2 - v_1 \\ v_3 - v_2 \\ \vdots \\ v_n - v_{n-1} \end{bmatrix}. \tag{2}$$

Z konstrukcji macierzy D wynika, że $D \in \mathbb{R}^{(n-1)\times n}$. Problem (1) jest bardzo trudnym zadaniem optymalizacji, możemy jednak, wykorzystując normę l_1, zastąpić je różnymi heurystykami, np.

$$\begin{aligned} &\underset{v}{\text{minimize}} && \|\tilde{y} - v\|_2^2 && (3a) \\ &\text{subject to} && \|Dv\|_1 \leqslant q && (3b) \end{aligned}$$

dla odpowiednio dobranej liczby q, lub

$$\underset{v}{\text{minimize}} \quad \|\tilde{y} - v\|_2^2 + \tau \|Dv\|_1 \tag{4}$$

dla odpowiednio dobranego parametru τ. Zadanie (4) spotyka się również pod nazwą LASSO (ang. *least absolute shrinkage and selection operator*). Sformułowanie LASSO jest jedną z podstawowych metod poszukiwania tzw. reprezentacji rzadkich, odgrywających ważną rolę między innymi w kompresji sygnałów [1]. Sformułowania (3) i (4) nie są sobie równoważne, nie są również równoważne zadaniu (1), będąc jedynie jego pewnym przybliżeniem. Niezwykle istotną cechą zadań (3) i (4) jest to, że są zadaniami wypukłymi. Estymatę $\hat{y}$ możemy też wyznaczyć rozwiązując modyfikacje (3) i (4) w postaci

$$\begin{aligned} &\underset{v}{\text{minimize}} && \|\tilde{y} - v\|_2 && (5a) \\ &\text{subject to} && \|Dv\|_1 \leqslant q && (5b) \end{aligned}$$

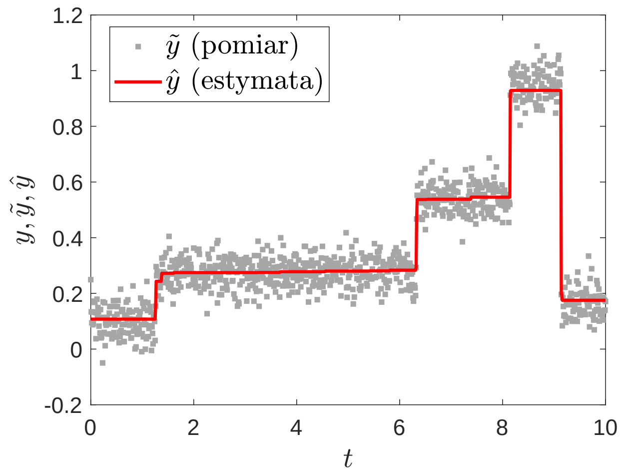
dla odpowiednio dobranej liczby q, oraz, odpowiednio,

$$\underset{v}{\text{minimize}} \quad \|\tilde{y} - v\|_2 + \tau \|Dv\|_1 \tag{6}$$

dla odpowiednio dobranego parametru τ.

2 Zadania

Zadanie 1. Pobrać plik danych `Data01.mat` (ISOD). Dla pobranych danych wyznaczyć rozwiązania zadań (3) i (4), dla różnych wartości parametrów q i τ. Należy skorzystać w tym celu z pakietu CVX [2]. Wygenerować wykresy podobne do przedstawionych na Rys. 1-2. Przy konstruowaniu macierzy D można skorzystać z funkcji `diag`. Należy zwrócić uwagę, że zamiast tworzyć macierz D, można zamiast Dv użyć składni `v(2:end)-v(1:end-1)`, gdzie `end` jest dostępnym w środowisku Matlab© obiektem, który zwraca wartość indeksu ostaniego elementu wektora.



Rysunek 2: Rekonstrukcja $\hat{y}$ sygnału kawałkami stałego y na podstawie pomiaru $\tilde{y}$.

Zadanie 2. Zamiast (5) i (6) możemy rozpatrzyć inne modyfikacje, np.

$$\begin{aligned} \underset{v}{\text{minimize}} \quad & \|\tilde{y} - v\|_1 && (7a) \\ \text{subject to} \quad & \|Dv\|_1 \leqslant q && (7b) \end{aligned}$$

dla odpowiednio dobranej liczby q, oraz, odpowiednio,

$$\underset{v}{\text{minimize}} \quad \|\tilde{y} - v\|_1 + \tau \|Dv\|_1 \tag{8}$$

dla odpowiednio dobranego parametru τ, które można sprawoadzić do postaci zadań LP [3] i rozwiązać np. korzystając z a) procedury `linprog`, b) procedury `solve` środowiska Matlab©. Sprowadzając (7) i (8) do zadań LP na ogół korzysta się z faktu, że zadanie optymalizacji

$$\begin{aligned} \underset{v}{\text{minimize}} \quad & f_0(v) && (9a) \\ \text{subject to} \quad & f_i(v) \leqslant 0, \quad i = 1, \ldots, m && (9b) \\ & h_i(v) = 0, \quad i = 1, \ldots, p && (9c) \end{aligned}$$

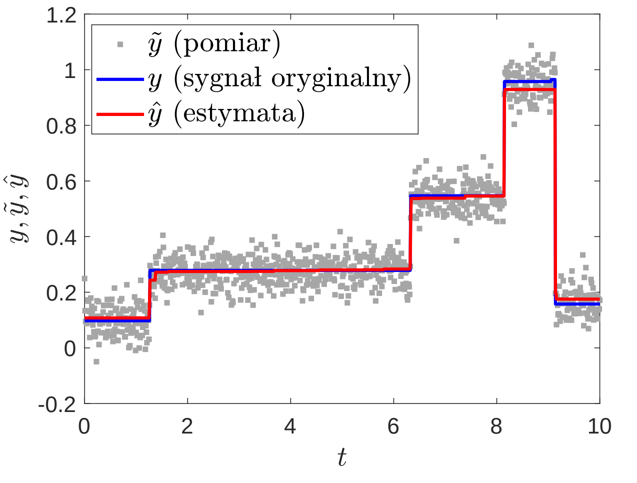
jest równoważne zadaniu

$$\begin{aligned} \underset{\varepsilon, v}{\text{minimize}} \quad & \varepsilon && (10a) \\ \text{subject to} \quad & f_0(v) \leqslant \varepsilon && (10b) \\ & f_i(v) \leqslant 0, \quad i = 1, \ldots, m && (10c) \\ & h_i(v) = 0, \quad i = 1, \ldots, p && (10d) \end{aligned}$$

Przekształć (7) i (8) do postaci LP

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^{\mathrm{T}} x && (11a) \\ \text{subject to} \quad & Ax \leqslant b && (11b) \end{aligned}$$

i wyznacz rozwiązania korzystając z a) procedury `linprog`, b) procedury `solve` środowiska Matlab©.



Rysunek 3: Rekonstrukcja sygnału kawałkami stałego. „Prawdziwy” sygnał y na ogól nie jest znany.

Wskazówka do Zadania 2

Rozpatrujemy zadanie

$$\begin{aligned} \underset{v}{\text{minimize}} \quad & \|\tilde{y} - v\|_1 && (12a) \\ \text{subject to} \quad & \|Dv\|_1 \leqslant q && (12b) \end{aligned}$$

gdzie $y \in \mathbb{R}^n$, $q \in \mathbb{R}$, $D \in \mathbb{R}^{m \times n}$ są dane, natomiast $v \in \mathbb{R}^n$ jest wektorem zmiennych optymalizacyjnych. Zadanie (12) jest równoważne zadaniu

$$\begin{aligned} \underset{v, \mu}{\text{minimize}} \quad & \sum_{i=1}^{n} |\tilde{y}_i - v_i| && (13a) \\ \text{subject to} \quad & \sum_{i=1}^{m} |\mu_i|_1 \leqslant q && (13b) \\ & \mu = Dv && (13c) \end{aligned}$$

gdzie

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix} \in \mathbb{R}^m. \tag{14}$$

Zadanie (13) jest równoważne zadaniu

$$\begin{aligned} \underset{v, \mu, \xi, \delta}{\text{minimize}} \quad & \sum_{i=1}^{n} \xi_i && (15a) \\ \text{subject to} \quad & |\tilde{y}_i - v_i| \leqslant \xi_i, \quad i = 1, \ldots, n && (15b) \\ & \sum_{i=1}^{m} \delta_i \leqslant q && (15c) \\ & |\mu_i| \leqslant \delta_i, \quad i = 1, \ldots, m && (15d) \\ & \mu = Dv && (15e) \end{aligned}$$

Zadanie (15) jest równoważne zadaniu

$$\begin{aligned} \underset{v, \mu, \xi, \delta}{\text{minimize}} \quad & \sum_{i=1}^{n} \xi_i && (16a) \\ \text{subject to} \quad & -\xi_i \leqslant \tilde{y}_i - v_i \leqslant \xi_i, \quad i = 1, \ldots, n && (16b) \\ & \sum_{i=1}^{m} \delta_i \leqslant q && (16c) \\ & -\delta_i \leqslant \mu_i \leqslant \delta_i, \quad i = 1, \ldots, m && (16d) \\ & \mu = Dv && (16e) \end{aligned}$$

Zadanie (16) jest równoważne zadaniu

$$\begin{aligned} \underset{v, \mu, \xi, \delta}{\text{minimize}} \quad & \mathbf{1}_n^{\mathrm{T}} \xi && (17a) \\ \text{subject to} \quad & -\xi \leqslant \tilde{y} - v \leqslant \xi && (17b) \\ & \mathbf{1}_m^{\mathrm{T}} \delta \leqslant q && (17c) \\ & -\delta \leqslant \mu \leqslant \delta && (17d) \\ & \mu = Dv && (17e) \end{aligned}$$

gdzie

$$\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{R}^n, \quad \delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_m \end{bmatrix} \in \mathbb{R}^m, \tag{18}$$

$$\mathbf{1}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n, \quad \mathbf{1}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^m. \tag{19}$$

Zadanie (17) jest równoważne zadaniu

$$\begin{aligned} \underset{v, \xi, \delta}{\text{minimize}} \quad & \mathbf{1}_n^{\mathrm{T}} \xi && (20a) \\ \text{subject to} \quad & -\xi \leqslant \tilde{y} - v \leqslant \xi && (20b) \\ & \mathbf{1}_m^{\mathrm{T}} \delta \leqslant q && (20c) \\ & -\delta \leqslant Dv \leqslant \delta && (20d) \end{aligned}$$

Zadanie (20) jest równoważne zadaniu

$$\begin{aligned} \underset{v,\xi,\delta}{\text{minimize}} \quad & \mathbf{1}_n^{\mathrm{T}}\xi && \text{(21a)} \\ \text{subject to} \quad & -\xi \leqslant \tilde{y} - v && \text{(21b)} \\ & \tilde{y} - v \leqslant \xi && \text{(21c)} \\ & \mathbf{1}_m^{\mathrm{T}}\delta \leqslant q && \text{(21d)} \\ & -\delta \leqslant Dv && \text{(21e)} \\ & Dv \leqslant \delta && \text{(21f)} \end{aligned}$$

czyli

$$\begin{aligned} \underset{v,\xi,\delta}{\text{minimize}} \quad & \mathbf{1}_n^{\mathrm{T}}\xi && \text{(22a)} \\ \text{subject to} \quad & \begin{bmatrix} I_n & -I_n & 0_{n\times m} \\ -I_n & -I_n & 0_{n\times m} \\ 0_{1\times n} & 0_{1\times n} & \mathbf{1}_m^{\mathrm{T}} \\ -D & 0_{m\times n} & -I_m \\ D & 0_{m\times n} & -I_m \end{bmatrix} \begin{bmatrix} v \\ \xi \\ \delta \end{bmatrix} \leqslant \begin{bmatrix} \tilde{y} \\ -\tilde{y} \\ q \\ 0_m \\ 0_m \end{bmatrix} && \text{(22b)} \end{aligned}$$

Przyjmując oznaczenia

$$A = \begin{bmatrix} I_n & -I_n & 0_{n\times m} \\ -I_n & -I_n & 0_{n\times m} \\ 0_{1\times n} & 0_{1\times n} & \mathbf{1}_m^{\mathrm{T}} \\ -D & 0_{m\times n} & -I_m \\ D & 0_{m\times n} & -I_m \end{bmatrix}, \quad c = \begin{bmatrix} 0_n \\ \mathbf{1}_n \\ 0_m \end{bmatrix}, \quad b = \begin{bmatrix} \tilde{y} \\ -\tilde{y} \\ q \\ 0_m \\ 0_m \end{bmatrix} \tag{23}$$

oraz

$$x = \begin{bmatrix} v \\ \xi \\ \delta \end{bmatrix}. \tag{24}$$

możemy zapisać zadanie (22) w postaci

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^{\mathrm{T}}x && \text{(25a)} \\ \text{subject to} \quad & Ax \leqslant b && \text{(25b)} \end{aligned}$$

Podobnie, można pokazać, że zadanie

$$\underset{v}{\text{minimize}} \quad \|\tilde{y} - v\|_1 + \tau \|Dv\|_1 \tag{26}$$

jest równoważne zadaniu

$$\begin{aligned} \underset{v,\xi,\delta}{\text{minimize}} \quad & \sum_{i=1}^{n} \xi_i + \tau \sum_{i=1}^{m} \delta_i && \text{(27a)} \\ \text{subject to} \quad & -\xi \leqslant x - v \leqslant \xi && \text{(27b)} \\ & -\delta \leqslant Dv \leqslant \delta && \text{(27c)} \end{aligned}$$

czyli

$$\begin{aligned} \underset{v,\xi,\delta}{\text{minimize}} \quad & \mathbf{1}_n^{\mathrm{T}}\xi + \tau \mathbf{1}_m^{\mathrm{T}}\delta && \text{(28a)} \\ \text{subject to} \quad & -\xi \leqslant x - v \leqslant \xi && \text{(28b)} \\ & -\delta \leqslant Dv \leqslant \delta && \text{(28c)} \end{aligned}$$

Sprowadzenie tego zadania do postaci (25) jest już relatywnie proste.

Literatura

[1] G.C. Calafiore and L. El Ghaoui. *Optimization Models*. Control systems and optimization series. Cambridge University Press, 2014.

[2] Inc. CVX Research. CVX: Matlab software for disciplined convex programming, version 2.0. `http://cvxr.com/cvx`, August 2012.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004. `http://web.stanford.edu/~boyd/cvxbook/`.