Dagstuhl 2014

Relation between total variation and persistence distance and its application in signal processing

Gerlind Plonka, Yi Zheng

University of Göttingen

February 2014

Outline

- Discrete total variation
- Definition of persistence distance
- Relation between TV and persistence distance
- Application in signal denoising
- Numerical experiments

Discrete Total Variation

Definition

Let $\mathbf{X}$ be a partition of the form $a = x_0 < x_1 < \cdots < x_n = b$ of the interval $[a, b]$. Let $\mathbf{y} = \{f(x_j)\}_{j=0}^{N}$ be a sequence corresponding to the partition $\mathbf{X}$.

Discrete total variation is defined as:

$$TV(\mathbf{y}) = TV(f) := \sum_{j=1}^{N} |f(x_j) - f(x_{j-1})|$$

Properties of discrete TV

Let $\mathbf{y} = \{f(x_j)\}_{j=0}^N$. Then

- $TV(\mathbf{y}) \geq 0$ and $TV(\mathbf{y}) = 0 \iff \mathbf{y} = c\mathbf{1}$.
- $TV(\lambda \mathbf{y}) = \lambda\, TV(\mathbf{y})$ for any $\lambda \geq 0$.
- $TV(\mathbf{y} + c\mathbf{1}) = TV(\mathbf{y})$.
- $TV(\mathbf{y}) : \mathbb{R}^{N+1} \to \mathbb{R}$ is a continuous functional.
- $TV(\mathbf{y}) + TV(\mathbf{z}) \geq TV(\max(\mathbf{y}, \mathbf{z})) + TV(\min(\mathbf{y}, \mathbf{z}))$, with $\max(\mathbf{y}, \mathbf{z}) := (\max\{y_j, z_j\})_{j=0}^N, \min(\mathbf{y}, \mathbf{z}) := (\min\{y_j, z_j\})_{j=0}^N$.
- $TV(\mathbf{y} + \mathbf{z}) \leq TV(\mathbf{y}) + TV(\mathbf{z})$.

Persistence Distance

1. Discrete TV sums up all absolute differences of neighboring function values (of a partition) without distinguishing the importance of those differences.
2. Homology Persistence provides a tool which can distinguish the importance of the changes of a function.

Persistence distance of a vector $\mathbf{y}$

Let $f \in S(\mathbf{X})$ with $\mathbf{X} : a = x_0 < x_1 < \ldots < x_n = b$ a linear spline.

Define

$$Y_m := \{y_k = f(x_k) : y_k \text{ is a local minimum value of } \mathbf{y}\},$$

$$Y^m := \{y_k = f(x_k) : y_k \text{ is a local maximum value of } \mathbf{y}\},$$

as well as the corresponding subsets of the partition $\mathbf{X}$,

$$X_m := \{x_k : f(x_k) \in Y_m\},$$

$$X^m := \{x_k : f(x_k) \in Y^m\}.$$

Algorithm to find persistence pairs

Input: Y_m, Y^m, X_m, X^m.

1. Set $P_1 := \emptyset$.
2. Fix $K_0 = \{f(x_{k_1}) \le f(x_{k_2}) \le \cdots \le f(x_{k_r})\}$ of maximum values in Y^m and the knot set $X_0 := X_m$.
3. For ℓ from 1 to r do
   Consider the ℓ-th entry $f(x_{k_\ell})$ in the ordered set K_0. If x_{k_ℓ} is not a boundary knot, then find the two special neighbors $\tilde{x}_1, \tilde{x}_2 \in X_{\ell-1}$ of x_{k_ℓ} and put

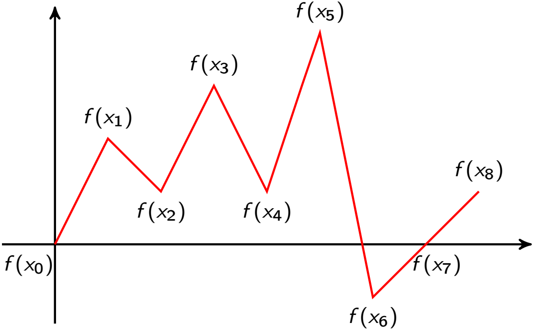
$$\tilde{x} := \operatorname*{argmin}_{x \in \{\tilde{x}_1, \tilde{x}_2\}} |\, f(x_{k_\ell}) - f(x) \,| \,.$$

   Then $(\tilde{x}, x_{k_\ell})$ is a persistence pair of f, and we set

$$P_1 = P_1 \cup (\tilde{x}, x_{k_l}),\ X_\ell := X_\ell \setminus \{\tilde{x}\}.$$

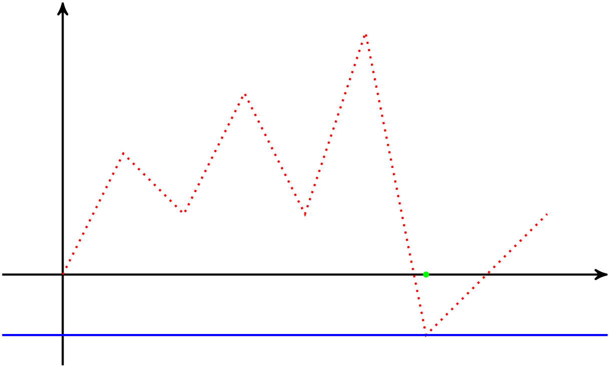
Output: P_1 containing all persistence pairs of $\mathbf{y}$ (resp. f).

Example



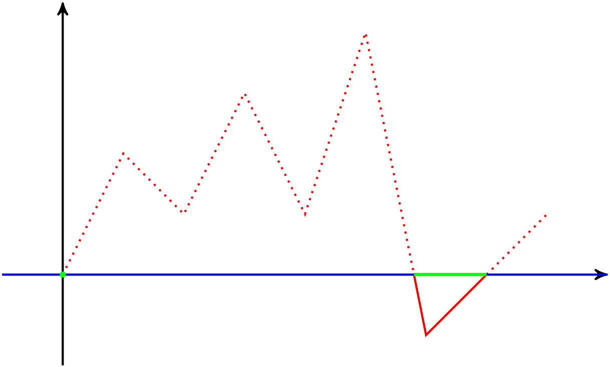
$$Y_m = \{f(x_0), f(x_2), f(x_4), f(x_6)\}, \qquad X_m = \{x_0,\, x_2,\, x_4,\, x_6\}$$
$$Y^m = \{f(x_1), f(x_3) f(x_5), f(x_8)\}\ , \qquad X^m = \{x_1,\, x_3,\, x_5,\, x_8\}$$

Example



blue: the line $y = t$
green: lower excursion set $f^{-1}((-\infty, t])$
$P_1 = \emptyset$

Example

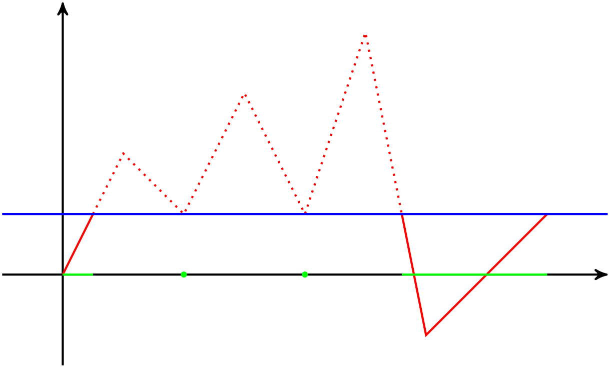


blue: the line $y = t$

green: lower excursion set $f^{-1}((-\infty, t])$

$P_1 = \emptyset$

Example

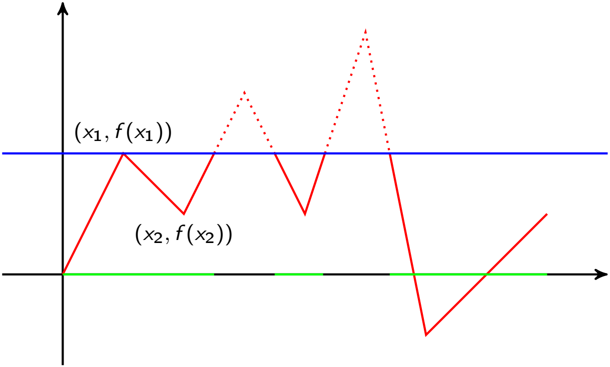


blue: the line $y = t$

green: lower excursion set $f^{-1}((-\infty, t])$

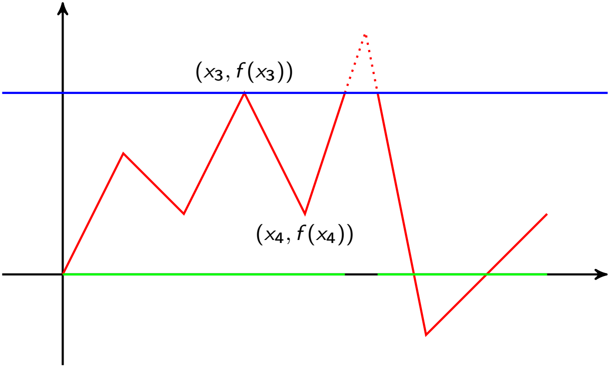
$P_1 = \emptyset$

Example



blue: the line $y = t$
green: lower excursion set $f^{-1}((-\infty, t])$
$P_1 = \{(x_1, x_2)\}$

Example

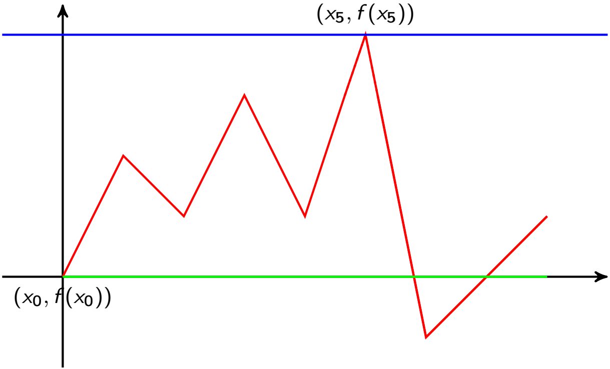


blue: the line $y = t$

green: lower excursion set $f^{-1}((-\infty, t])$

$P_1 = \{(x_1, x_2), (x_3, x_4)\}$

Example



blue: the line $y = t$

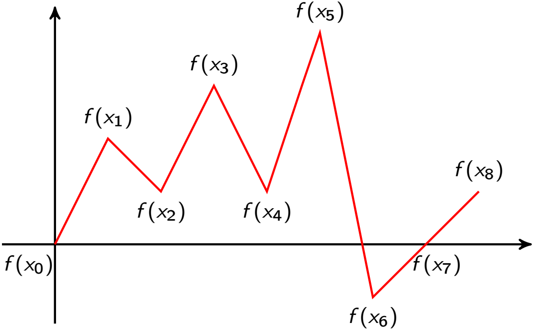
green: lower excursion set $f^{-1}((-\infty, t])$

$P_1 = \{(x_1, x_2), (x_3, x_4), (x_0, x_5)\}$

Persistence distance of a vector $\mathbf{y}$

- We get P_1, the set of all persistence pairs obtained by going from below to above.
- We apply the same rule from above to below and get the set of persistence pairs P_2.

Example

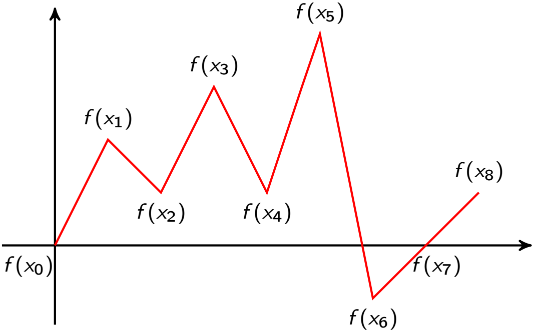


$$P_1 = \{(x_1, x_2), (x_3, x_4), (x_0, x_5)\}$$
$$P_2 = \{(x_1, x_2), (x_3, x_4), (x_6, x_8)\}$$

Persistence distance of a vector $\mathbf{y}$

- We get P_1, the set of all persistence pairs which are obtained by going from below to above.
- We apply the same rule from above to below and get the set of persistence pairs P_2.
- Usually, $P_1 \cap P_2 \neq \emptyset$ and $P_1 \neq P_2$.

Example



$P_1 = \{(x_1, x_2), (x_3, x_4), (x_0, x_5)\}$,
Consider $|f(x_1) - f(x_2)| + |f(x_3) - f(x_4)| + |f(x_0) - f(x_5)|$
$P_2 = \{(x_1, x_2), (x_3, x_4), (x_6, x_8)\}$,
Consider $|f(x_1) - f(x_2)| + |f(x_3) - f(x_4)| + |f(x_6) - f(x_8)|$

Persistence Distance

Let $S_1(\mathbf{X})$ be the space of linear spline functions corresponding to the partition $\mathbf{X}$ and $\mathbf{y} = (f(x_k))_{k=0}^n$.

Definition (Plonka, Zheng)

For a given function $f \in S_1(\mathbf{X})$, the **Persistence Distance** of f resp. $\mathbf{y}$ is defined as

$$\|\mathbf{y}\|_{per} = \|f\|_{per} := \sum_{(x_k,x_\ell)\in P_1} |f(x_\ell) - f(x_k)| + \sum_{(x_k,x_\ell)\in P_2} |f(x_\ell) - f(x_k)|,$$

i.e., as the sum of all distances of function values for the persistence pairs in P_1 and P_2.

Properties of Persistence Distance

Theorem [Plonka, Zheng (2013)]

(i) $\|\mathbf{y}\|_{per} \geq 0$. $\|\mathbf{y}\|_{per} = 0 \iff (y_j)_{j=0}^{n}$ is a monotone sequence.

(ii) For $c \in \mathbb{R}$, $\|c\mathbf{y}\|_{per} = |c| \cdot \|\mathbf{y}\|_{per}$.

(iii) $\|\mathbf{y} + c\mathbf{1}\|_{per} = \|\mathbf{y}\|_{per}$,

(iv) $\|\mathbf{y}\|_{per} : S_1(\mathbf{X}) \to \mathbf{R}$ is a lower semi-continuous functional.

(v) $\|\mathbf{y}\|_{per} + \|\mathbf{z}\|_{per} \geq \|\max(\mathbf{y},\mathbf{z})\|_{per} + \|\min(\mathbf{y},\mathbf{z})\|_{per}$, where

$$\max(\mathbf{y},\, \mathbf{z}) := (\max\{y_j, z_j\})_{j=0}^{N}, \min(\mathbf{y},\, \mathbf{z}) := (\min\{y_j, z_j\})_{j=0}^{N}.$$

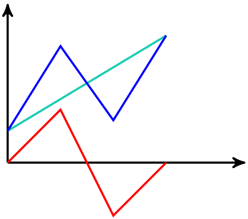
Properties of Persistence Distance

(vi) The persistence distance $\|\mathbf{y}\|_{per}$ does not satisfy the triangle inequality

$$\|\mathbf{y}+\mathbf{z}\|_{per} \nleq \|\mathbf{y}\|_{per} + \|\mathbf{z}\|_{per},$$

and hence, is not convex.

Counter example for (vi)



$f = (0, 1, -1, 0)$, $P_1 = \{(x_0, x_1)\}$, $P_2 = \{(x_3, x_2)\}$, $\|f\|_{per} = 2$.

$g = (0.6, 1.2, 1.8, 2.4)$, $P_1 = P_2 = \emptyset$, $\|g\|_{per} = 0$.

$f + g = (0.6, 2.2, 0.8, 2.4)$, $P_1 = \{(x_2, x_1)\}$, $P_2 = \{(x_1, x_2)\}$, $\|f + g\|_{per} = 2.8$.

$\|f + g\|_{per} > \|f\|_{per} + \|g\|_{per}$

Relation between discrete TV and persistence distance

Theorem (Plonka, Zheng '13)

Let $\mathbf{X}$ *be a partition of the form* $a = x_0 < x_1 < \cdots x_N = b$. *Then, for each function* $f \in S_1(\mathbf{X})$ *we have*

$$\|f\|_{per} + \max_{x,y \in \mathbf{X}} |f(x) - f(y)| = TV(f)$$

Analogously, for each sequence $\mathbf{y} \in \boldsymbol{R}^{N+1}$, *we have*

$$\|\mathbf{y}\|_{per} + \max_{j,k \in \{0,1,\ldots,N\}} |y_j - y_k| = TV(\mathbf{y}).$$

Application to signal denoising

- Discrete ROF model for signal denoising

$$\min J(u) = \frac{\lambda}{2} \sum_{k=0}^{N} |u(x_k) - f(x_k)| + TV(u)$$

Idea: Replace $TV(u)$ by

$$TV(u) = \|u\|_{per} + \max_{j,k \in \{0,1,\ldots,N\}} |u(x_j) - u(x_k)|$$

- Apply different weights for different pairs.

Application to signal denoising

Consider a weighted ROF-model based on persistence distance

$$\tilde{J}(u) = \frac{\lambda}{2} \sum_{j=0}^{N} |u(x_j) - f(x_j)|^2 + \sum_{(x_j, \tilde{x}_j) \in P(u)} \alpha_j(u) |u(x_j) - u(\tilde{x}_j)|,$$

Choose the weights

$$\alpha_j(u) = \frac{1}{1 + \beta |u(\tilde{x}_j) - u(x_j)|}, \qquad \beta > 0.$$

Application to signal denoising

Theorem (Plonka, Zheng '13)

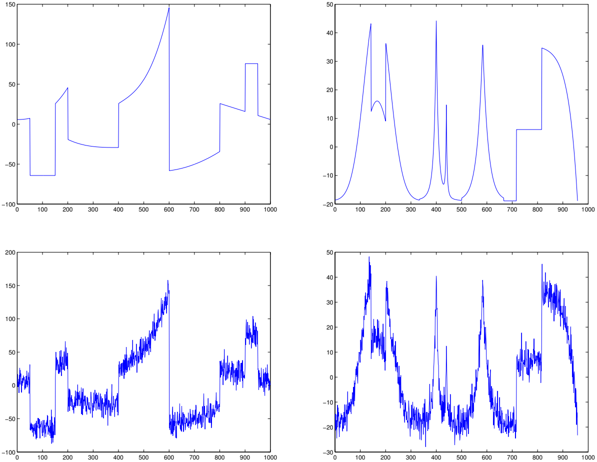
Consider for each $[x_\ell, x_{\ell+1}]$, $\ell = 0, \ldots, N-1$, *the complete chain of persistence intervals* $[x_\ell, x_{\ell+1}] \subseteq [x_1^\ell, \tilde{x}_1^\ell] \subset \ldots \subset [x_{r(\ell)}^\ell, \tilde{x}_{r(\ell)}^\ell]$.
Denote by $\alpha_\nu^\ell(u)$ *the weight in* $\tilde{J}(u)$ *corresponding to the persistence pair* $(x_\nu^\ell, \tilde{x}_\nu^\ell)$. *Then the weighted functional* $\tilde{J}(u)$ *is equivalent to*

$$J_w(u) := \frac{\lambda}{2} \sum_{j=0}^{N} |u(x_j) - f(x_j)|^2 + \sum_{\ell=0}^{N-1} w_\ell(u) |u(x_{\ell+1}) - u(x_\ell)|,$$

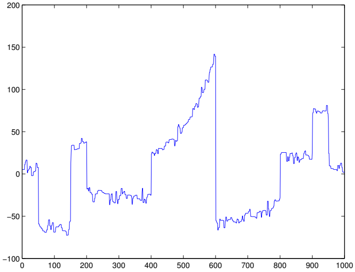
where

$$w_\ell(u) = w_l := \sum_{\nu=1}^{r(\ell)} (-1)^{\nu-1} \alpha_\nu^\ell.$$

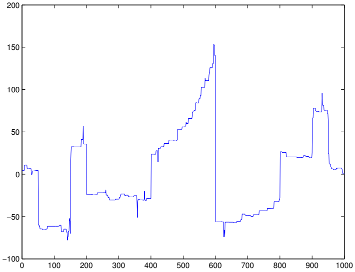
Application to signal denoising



Application to signal denoising

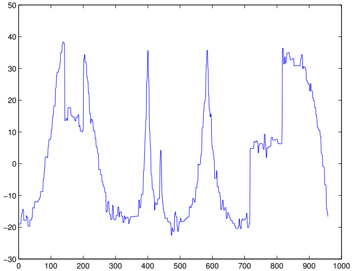


(a) COS method, PSNR 33.79

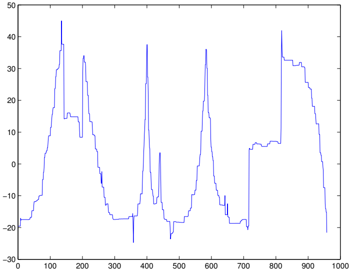


(b) persistence denoising, PSNR 35.56

Application to signal denoising



(a) COS method, PSNR 31.21



(b) persistence denoising, PSNR 31.80

Thank you

Thank you for your attention.