The calculus of p-variation

June 14, 2014

- $PP[a,b]$ – the set of all point partitions of $[a,b]$ (def. 1.1).
- $s_p(f;\kappa)$ – p-variation sum (def. 1.2).
- $v_p\,(f)$ – p-variation of the function f (def. 1.2).
- $MP_p(f,[a,b])$ – the set of meaningful partitions (def. 1.2).
- $PM[a,b]$ – a set of piecewise monotone functions (def. 1.6).
- $CPM[a,b]$ – a set of continuous piecewise monotone functions (def. 1.6).
- $K(f,[a,b])$ – minimal size of PM partitions (prop. 1.7).
- $X(f,[a,b])$ – the set of PM partitions with minimal size (def. 1.8).
- $\overline{BP}(f,[a,b])$ – the set of break points (definition 1.25).
- $\overline{BP}(f,(a,b))$ – the set of inner break points (def. 1.25).
- $BP(f,[a,b])$ – the set of partitions that contains only break points (def. 1.25).
- $BP(f,(a,b))$ – the set of partitions that contains only break points, but excludes the partition $\{a,b\}$. (def. 1.25).
- $\overline{MB}_p(f,[a,b])$ – the set of meaningfully break points (def. 1.27).
- $MB_p(f,[a,b])$ – the set of partitions that contains only meaningfully break points (def. 1.27).

1 Mathematical analysis

1.1 General known properties

Definition 1.1 (Partition). Let $J = [a, b]$ be a closed interval of real numbers with $-\infty < a \leq b < +\infty$. If $a < b$, an ordered set $\kappa = \{x_i\}_{i=0}^n$ of points in $[a, b]$ such that a $a = x_0 < x_1 < x_2 < ... < x_n = b$ is called a *(point) partition.* The size of the partition is denoted $|\kappa| := \#\kappa - 1 = n$. The set of all point partitions of $[a, b]$ is denoted by $PP[a, b]$.

Definition 1.2 (p-variation). Let $f : [a, b] \to \mathbb{R}$ be a real function from an interval $[a, b]$. If $a < b$, for $\kappa = \{x_i\}_{i=0}^n \in PP[a, b]$ the *p-variation sum* is

$$s_p(f, \kappa) := \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|^p, \tag{1.1}$$

where $0 < p < \infty$. Thus, the *p-variation* of f over $[a, b]$ is 0 if $a = b$ and otherwise

$$v_p(f) = v_p(f, [a, b]) := \sup \{s_p(f, \kappa) : \kappa \in PP[a, b]\}. \tag{1.2}$$

The partition κ is called *supreme partition* if it satisfies the property $v_p(f) = s_p(f, \kappa)$. The set of such partitions is denoted $SP_p(f, [a, b])$.

Lemma 1.3 (Elementary properties). Let $f : [a, b] \to \mathbb{R}$ and $0 < p < \infty$. Then the following p-variation properties holds

a) $v_p(f, [a, b]) \geq 0$,

b) $v_p(f, [a, b]) = 0 \Leftrightarrow f \equiv Const.$,

c) $\forall C \in \mathbb{R} : v_p(f + C, [a, b]) = v_p(f, [a, b])$,

d) $\forall C \in \mathbb{R} : v_p(Cf, [a, b]) = C^p v_p(f, [a, b])$,

e) $\forall c \in [a, b] : v_p(f, [a, b]) \geq v_p(f, [a, c]) + v_p(f, [c, b])$,

f) $\forall [a', b'] \subset [a, b] : v_p(f, [a, b]) \geq v_p(f, [a', b'])$.

g) $\forall \kappa \in PP[a, b] : s_p(f; \kappa) \leq v_p(f, [a, b])$.

All listed properties are elementary derived directly form the p-variation definition.

Definition 1.4 (Regulated function). ([1], Def. 3.1) For any interval J, which may be open or closed at either end, real function f is called *regulated* on J if it has left and right limits $f(x-)$ and $f(x+)$ respectively at each point x in interior of J, a right limit at the left end point and a left limit at the right endpoint.

Proposition 1.5. ([1], Lemma 3.1) Let $1 \leq p < \infty$. If f is regulated then $v_p(f)$ remains the same if points $x+$, $x-$ are allowed as partition points x_i in the definition 1.2.

Definition 1.6 (Piecewise monotone functions). ([1], Def. 3.2) A regulated real-valued function f on closed interval $[a, b]$ will be called *piecewise monotone* (PM) if there are points $a = x_0 < \cdots < x_k = b$ for some finite k such that f is monotone on each interval $[x_{j-1}, x_j]$, $j = 1, \dots, k$. Here for $j = 1, \dots, k-1$, x_j may be point $x-$ or $x+$. The set of all piecewise monotone functions is denoted $PM = PM[a, b]$.

In addition to PM, if f is continuous function we will call it continuous piecewise monotone(CPM). The set of such functions is denoted $CPM = CPM[a, b]$.

Proposition 1.7. ([1], Prop. 3.1) If f is PM, there is a minimal size of partition $|\kappa|$ for which the definition 1.6 holds. The minimal size of the the PM partition is denoted $K(f, [a, b]) = K(f)$, namely

$$K(f) := \min \left\{ n : \exists \{x_i\}_{i=0}^{n} \in PP[a, b] : f \text{ is monotonic in each } [x_{j-1}, x_j] \right\}. \tag{1.3}$$

Definition 1.8 (The set of PM partitions with minimal size). ([1], Def. 3.3) If f is PM, let $X(f) = X(f, [a, b])$ be the set of all $\{x_i\}_{i=0}^{K(f)}$ for which the definition of PM (def. 1.6) holds. $X(f)$ is called the *set of PM partitions with minimal size.*

Proposition 1.9. ([1], Prop. 3.3) Let f is PM then the numbers $\alpha_j(f) := f(x_j) - f(x_{j-1})$ for $\{x_j\}_{j=0}^{K(f)} \in X(f)$ and $j = 1, 2, \dots K(f)$ are uniquely determined.

Definition 1.10 (The equality by PM). ([1], Def. 3.4) If f, g are two PM functions, possibly on different intervals, such that $K(f) = K(g)$ and $\alpha_j(f) = \alpha_j(g)$ for $j = 1, 2, \dots, K(f)$, then we say that $f \overset{PM}{=} g$.

Proposition 1.11. ([1], Cor. 3.1) Let $p > 1$ and functions f and g are PM. If $f \overset{PM}{=} g$ or $f \overset{PM}{=} -g$, then $v_p(f) = v_p(g)$.

Proposition 1.12. ([1], Them. 3.1) Let f is PM, $\kappa \in X(f)$ and $1 \leq p < \infty$. Then the supremum of p-variation in Definition 1.2 is attained for some partition $r \subset \kappa$.

Corollary 1.13. The set $SP_p(f, [a, b])$ is not empty for all $f \in PM[a, b]$.

Definition 1.14 (Sample function). Suppose $X = \{X_i\}_{i=0}^n$ is any sequence real numbers. We will call such sequence a *sample*, whereas n will be referred to as a *sample size*. Then the *sample function* $G_X : [0, n] \to \mathbb{R}$ is defined as

$$G_X(t) := X_{\lfloor t \rfloor},\ t \in [0, n], \tag{1.4}$$

where $\lfloor t \rfloor$ denotes floor function at point t.

Definition 1.15 (p-variation of the sequence). Let $X = \{X_i\}_{i=0}^n$. The p-variation of the sample X is defined as p-variation of the function $G_X(t)$, namely

$$v_p(X) := v_p(G_X(t), [0, n]). \tag{1.5}$$

Definition 1.16 (Piecewise linear function). Let $X = \{X_i\}_{i=0}^n$ be any real-value sequence. The function $L_X : [0, n] \to \mathbb{R}$ is defined as

$$L_X(t) := (1 + \lfloor t \rfloor - t)X_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)X_{\lfloor t \rfloor + 1},\ t \in [0, n] \tag{1.6}$$

is called *piecewise linear function*.

Definition 1.17 (Partial sum). Let $X_1, X_2, \ldots, X_n$ be any sequence of real numbers. The *partial sum* of the first j terms is defined by

$$S_j := \sum_{i=1}^{j} X_i,\ j = 1, 2, \ldots, n. \tag{1.7}$$

In addition, lets denote $S_0 = 0$.

1.2 General properties with proofs

Proposition 1.18. For all $f \in PM$ exists $g \in CPM$ such that $f \overset{PM}{=} g$.

Proof. Let f be PM. According to Proposition 1.9 the numbers $\alpha_j(f) := f(x_j) - f(x_{j-1})$, $j = 1, 2, \ldots K(f)$ are uniquely determined. Then, the partial sums of the sequence $\alpha_j(f)$ are

$$S_j := \sum_{i=1}^{j} \alpha_i(f). \tag{1.8}$$

Lets connect points S_j by piecewise linear function (def. 1.16), namely

$$L_S(t) := (1 + \lfloor t \rfloor - t)S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)S_{\lfloor t \rfloor + 1},\ t \in [0, K(f)] \tag{1.9}$$

Function L_S is CPM. In addition, it is straight forward to see that

$$\alpha_j(L_S) = S_j - S_{j-1} = \alpha_j(f), \tag{1.10}$$

hence, by Definition 1.10 $f \overset{PM}{=} L_S$. ■

Corollary 1.19. By applying last proof to the function $G_X(t)$ (from Definition 1.14), we get that

$$G_X(t) \overset{PM}{=} L_X(t). \tag{1.11}$$

Therefore, according to Proposition 1.11, if $p > 1$ then

$$v_p(X) := v_p(G_X(t), [0, n]) = v_p(L_X(t), [0, n]) \tag{1.12}$$

Proposition 1.20. Suppose $f : [a, b] \to \mathbb{R}$ is CPM. Let $\{x_i\}_{i=0}^n \in PP[a, b]$ be any partition of interval $[a, b]$. Then the statement

$$\exists \kappa \in SP_p(f, [a, b]) : \forall i, x_i \in \kappa \tag{1.13}$$

is equivalent to

$$v_p(f, [a, b]) = \sum_{i=1}^{n} v_p(f, [x_{i-1}, x_i]). \tag{1.14}$$

Proof. Necessary. Let $f : [a, b] \to \mathbb{R}$ be CPM, $\{x_i\}_{i=0}^n \in PP[a, b]$ and

$$\exists \kappa \in SP_p(f, [a, b]) : \forall i, x_i \in \kappa. \tag{1.15}$$

Points from the partition κ will be denoted t_i, i.e. $\kappa = \{t_i\}_{i=0}^m$. Then, according to definitions of SP_p and p-variation (def. 1.2) the following equation holds

$$v_p(f, [a, b]) = s_p(f; \kappa) = \sum_{j=1}^{m} |f(t_{j-1}) - f(t_j)|^p = \sum_{i=1}^{n} \sum_{j=h(i-1)+1}^{h(i)} |f(t_{j-1}) - f(t_j)|^p, \tag{1.16}$$

where $h : \{0, 1, \ldots, n\} \to \{0, 1, \ldots, m\}$ denotes a function from the set of index of x to the set of index of t, namely:

$$h(i) := (j_i : x_i = t_{j_i} = t_{h(i)}). \tag{1.17}$$

The equation (1.16) holds, because all the elements in the sum remains, we just grouped them.

Moreover, the inequality

$$\sum_{j=h(i-1)+1}^{h(i)} |f(t_{j-1}) - f(t_j)|^p \leq v_p(f, [x_{i-1}, x_i]) \tag{1.18}$$

holds according to Lemma 1.3(g).

As a result of (1.16) and (1.18) we get

$$v_p(f, [a, b]) \leq \sum_{i=1}^{n} v_p(f, [x_{i-1}, x_i]). \tag{1.19}$$

On the other hand, according to the same Lemma 1.3(e) the following inequality holds

$$v_p(f, [a, b]) \geq \sum_{i=1}^{n} v_p(f, [x_{i-1}, x_i]). \tag{1.20}$$

Finally, from the (1.19) and (1.20) follows

$$v_p(f, [a, b]) = \sum_{i=1}^{n} v_p(f, [x_{i-1}, x_i]). \tag{1.21}$$

Sufficiency. Suppose $f : [a, b] \to \mathbb{R}$ is CPM and

$$v_p(f, [a, b]) = \sum_{i=1}^{n} v_p(f, [x_{i-1}, x_i]). \tag{1.22}$$

According to Corollary 1.13, sets $SP_p(f, [x_{i-1}, x_i]),\ i = 1, \ldots, n$ are not empty. Lets take any partition from each of the sets $SP_p(f, [x_{i-1}, x_i]),\ i = 1, \ldots, n$ and denote it κ_i.

Then, let define a joint partition $\kappa := \cup_{i=1}^{n} \kappa_i$. Points from the partition κ will be denoted by t_i. In addition, we will use the function h, which is defined in (1.17). Then, continuing the equation (1.22) we get

$$v_p(f, [a, b]) = \sum_{i=1}^{n} \sum_{j=h(i-1)+1}^{h(i)} |f(t_{j-1}) - f(t_j)|^p = \sum_{i=1}^{|\kappa|} |f(t_{i-1}) - f(t_i)|^p. \tag{1.23}$$

This means that $\kappa \in MP_p(f, [a, b])$. Moreover, $\forall i : x_i \in \kappa$, because $\kappa = \cup_{i=1}^{n} \kappa_i$. ■

1.3 Meaningful break points

Definition 1.21 (Lexicographical order). Suppose f is CPM. Let κ and r denotes two partitions with minimal size, namely $\kappa = \{x_j\}_{j=0}^{K(f)} \in X(f)$ and $r = \{y_j\}_{j=0}^{K(f)} \in X(f)$. Then partitions κ and r could be compare using *lexicographical order*, namely we will say that $\kappa \succ r$, if

- $\exists j : x_j > y_j$; and
- $\forall i < j : x_i = y_i$.

We will say $\kappa = r$ if $\forall j : x_j = y_j$.
We will say $\kappa \succeq r$ if $\kappa \succ r$ or $\kappa = r$.

Proposition 1.22. Binary relation $\succeq$ is a total order, i.e. the relationship $\succeq$ satisfies the flowing properties for all $\kappa, r, g \in X(f)$:

- Antisymmetry – if $\kappa \succeq r$ and $r \succeq \kappa$, then $\kappa = r$.
- Transitivity – if $\kappa \succeq r$ and $r \succeq g$, then $\kappa \succeq g$.
- Totality – either $\kappa \succeq r$ or $r \succeq \kappa$.

Proof. Need reference ... ■

Proposition 1.23. Let $f : [a,b] \to \mathbb{R}$ is CPM. Then there is a partition $\kappa_f \in X(f)$ such that

$$\forall r \in X(f) : \kappa_f \succeq r. \tag{1.24}$$

Proof. Firstly, we will construct κ_f, latter, we will proof that the property (1.24) holds.

Suppose, $f : [a,b] \to \mathbb{R}$ is CPM. Let denote

$$x_1 := \sup \left\{ x \in [a,b] : |f(y_1) - f(a)| \leq |f(y_2) - f(a)|, \forall a \leq y_1 < y_2 \leq x \right\}.$$

Other $x_i,\ i = 2, \dots, K(f)$ will be defined by induction. Suppose x_{i-1} is defined and $x_{i-1} < b$, then

$$x_i := \sup \left\{ x \in [x_{i-1}, b] : |f(y_1) - f(x_{i-1})| \leq |f(y_2) - f(x_{i-1})|, \forall x_{i-1} \leq y_1 < y_2 \leq x \right\}.$$

According condition used in the definition of x_i, all $(x_{i-1}, x_i),\ i = 1, \dots, K(f)$ are monotonic intervals.

In addition, points x_i have the greatest values, that satisfies the condition. Therefore, if the values x_i would increased, the condition would no longer be valid, i.e.

$$\forall i, \exists \delta : |f(x_i) - f(x_{i-1})| > |f(x_i + \varepsilon) - f(x_{i-1})|, \text{ if } \varepsilon \in (0, \delta). \tag{1.25}$$

We will show that $\kappa_f := \{x_i\}_{i=1}^n$ satisfies the condition (1.24). Suppose to the contrary that

$$\exists r = \{t_i\}_{i=1}^n \in X(f) : r \succ \kappa_s. \tag{1.26}$$

According to Definition of $\succ$ we get

$$\exists j : t_j > x_j, \text{ and } t_i = x_i, \text{ if } i < j. \tag{1.27}$$

Since $r \in X(f)$, the intervals $[t_{j-1}, t_j]$ are monotonic intervals, i.e.

$$|f(y_1) - f(t_{j-1})| \leq |f(y_2) - f(t_{j-1})|, \tag{1.28}$$

if $t_{j-1} \leq y_1 < y_2 \leq t_j$. But (1.28) contradicts (1.25), then $y_1 = x_j$ and $y_2 = x_j + \varepsilon$. ∎

Lemma 1.24. Suppose $f : [a, b] \to \mathbb{R}$ is CPM and $\kappa_f = \{x_i\}_{i=1}^{K(f)}$ is the partition from (1.24). Then, for all $x_i, i = 1, \dots, K(f) - 1$, the following proposition holds

$$\forall \varepsilon > 0 : f \text{ is not a constant in } [x_i, x_i + \varepsilon]. \tag{1.29}$$

Proof. Suppose $f : [a, b] \to \mathbb{R}$ is CPM and $\kappa_f = \{x_i\}_{i=1}^{K(f)}$ is the partition from (1.24). Suppose to the contrary, that exists x_j and $\varepsilon > 0$ such that f is not a constant in $[x_j, x_j + \varepsilon]$. Since f is constant in $[x_j, x_j + \varepsilon]$, the point $x_j + \varepsilon$ is also the end of monotone interval (def. 1.6).

Lets define the partition $\{y_i\}_{i=1}^{K(f)} \in X(f)$, which is

$$y_i := \begin{cases} x_i & , \text{if } j \neq i \\ x_i + \varepsilon & , \text{if } j = i \end{cases}, \tag{1.30}$$

where $\varepsilon \in (0, x_{j+1} - x_j)$.

It is straight forward to see that $\{y_i\}_{i=1}^{K(f)} \succ \kappa_f$, since $y_j > x_j$ and $y_i = x_i, i < j$. This contradict the definition of κ_f. ∎

Definition 1.25 (Break points). Let κ_f be a partition form Proposition 1.23. The points $x_j \in \kappa_f$ will be called *break points* and the set of such points are denoted $\overline{BP}(f, [a, b])$, which is non empty and finite. The set of partitions that contains only break points are denoted $BP(f, [a, b])$, i.e.

$$BP(f) = BP(f, [a, b]) := \left\{\kappa \in PP[a, b] : \kappa \subset \overline{BP}(f, [a, b])\right\}. \tag{1.31}$$

$BP(f, [a, b])$ is also non empty and finite.

For the convenience let define the subsets of the sets $\overline{BP}$ and BP that excludes the ends of the interval, namely

$$\overline{BP}(f, (a, b)) := \overline{BP}(f, [a, b]) \setminus \{a, b\}, \tag{1.32}$$
$$BP(f, (a, b)) := BP(f, [a, b]) \setminus \{\{a, b\}\}. \tag{1.33}$$

Proposition 1.26. Let f is CPM function defined in $[a, b]$ and $1 \leq p < \infty$. Then the p-variations could be expressed as

$$v_p(f) = v_p(f, [a, b]) = \max \{s_p(f; \kappa) : \kappa \in BP(f, [a, b])\}. \quad (1.34)$$

Proof. This result follows directly from the Proposition 1.12. ■

Definition 1.27 (Meaningfull break points). Suppose f is CPM and $1 \leq p < \infty$. Let denote the set of *partitions of meaningful break* (points)

$$MB_p(f) = MB_p(f, [a, b]) := MP_p(f, [a, b]) \cap BP(f, [a, b]). \quad (1.35)$$

The set $MB_p(f, [a, b])$ is not empty according to Proposition 1.26.

The point x will be called *meaningful break* if $\exists \kappa \in MB_p(f) : x \in \kappa$. The set of such points are denoted $\overline{MB}_p(f) = \overline{MB}_p(f, [a, b])$. The point a is called *meaningless* if $a \notin \overline{MB}_p(f, [a, b])$.

Lemma 1.28. Let $\{x_i\}_{i=0}^n \subset \kappa \in MB_p(f, [a, b])$, then

$$\forall i, j : v_p(f, [x_i, x_j]) = \sum_{k=i+1}^{j} v_p(f, [x_{k-1}, x_k]),\ 0 \leq i < j \leq n. \quad (1.36)$$

Proof. Suppose $\{x_i\}_{i=0}^n \subset \kappa \in MB_p(f, [a, b])$. Let choose i and j such that $0 \leq i < j \leq n$. Because $MB_p \subset MP_p$, we can apply Proposition 1.20 for the partition $\{x_0, x_i, x_{i+1}, \ldots, x_{j-1}, x_j, x_n\}$. Thus,

$$v_p(f, [a, b]) = v_p(f, [x_0, x_i]) + \sum_{k=i+1}^{j} v_p(f, [x_{k-1}, x_k]) + v_p(f, [x_j, x_n]). \quad (1.37)$$

In addition, we can apply the same proposition for the partition $\{x_0, x_i, x_j, x_n\}$, then

$$v_p(f, [a, b]) = v_p(f, [x_0, x_i]) + v_p(f, [x_i, x_j]) + v_p(f, [x_j, x_n]). \quad (1.38)$$

By subtracting one equation form the other we get the result that

$$v_p(f, [x_i, x_j]) = \sum_{k=i+1}^{j} v_p(f, [x_{k-1}, x_k]). \quad (1.39)$$

■

Proposition 1.29. Suppose $f : [a, b] \to \mathbb{R}$ is CPM. Then

$$\exists \kappa_{ab} \in MB_p(f, [a, b]) : \forall \kappa \in MB_p(f, [a, b])[|\kappa_{ab}| \geq |\kappa|]. \quad (1.40)$$

Proof. Suppose f is CPM, then $\forall \kappa \in BP(f) : |\kappa| \leq K(f)$. Since $MB_p(f) \subset BP(f)$, then $\forall \kappa \in MP_p(f) : |\kappa| \leq K(f)$. Therefore, the size of partitions from $MP_p(f)$ are bounded. As a result the size of the partition $|\kappa|$ is bounded natural number, whereas such set has the larges element. ■

Definition 1.30. The set of partitions κ_{ab} which satisfies 1.40 will be denoted $MB_p^m(f, [a, b]) = MB_p^m(f)$. This set is non empty according to Proposition 1.29.

In addition it could be shown (see Proposition 1.45) that the set $MB_p^m(f)$ has exactly one element and

$$\forall \kappa \in MB_p(f) : \kappa \subset \kappa_{ab} \in MB_p^m(f) \tag{1.41}$$

Lemma 1.31. Let f is CMD. Then $\{a, b\} \in MB_p^m(f, [a, b])$ if and only if

$$\forall \kappa \in BP(f, (a, b)) : v_p(f, [a, b]) > s_p(f, \kappa).$$

Here the $BP(f, (a, b))$ is the set of partitions of break points excluding $\{a, b\}$ (see def. 1.33).

Proof. Necessary. Let assume on the contrary that $\{a, b\} \in MP_p^m(f, [a, b])$, but

$$\exists \kappa \in BP(f, (a, b)) : v_p(f, [a, b]) = s_p(f, \kappa).$$

The partition $\kappa \in BP(f, (a, b))$ must have more points then the end of the intervals a and b, because if $\kappa = \{a, b\}$ then it contradicts to the definition of $BP(f, (a, b))$. Therefore, $|\kappa| > |\{a, b\}|$, but this contradicts to assumptions that $\{a, b\} \in MP_p^m(f, [a, b])$.

Sufficiency. Suppose

$$\forall \kappa \in BP(f, (a, b)) : v_p(f, [a, b]) > s_p(f, \kappa). \tag{1.42}$$

According to Proposition 1.26, the p-variation is achieved in the set $BP(f, [a, b])$, therefore, from the (1.42) follows that $\{a, b\}$ is the partition of meaningful break points, i.e. $\{a, b\} \in MP_p(f, [a, b])$. In addition, this partition is the only one, therefore it is the biggest partition, thus $\{a, b\} \in MP_p^m(f, [a, b])$. ■

Lemma 1.32. Let $f : [a, b] \to \mathbb{R}$ be CPM. Suppose $\{x_i\}_{i=0}^n \in MB_p^m(f, [a, b])$. Then

$$\forall i \in \{1, 2, \ldots, n\} : \{x_{i-1}, x_i\} \in MB_p^m(f, [x_{i-1}, x_i]). \tag{1.43}$$

Proof. Suppose to the contrary that

$$\exists j \in \{1, 2, \ldots, n\} : \{x_{j-1}, x_j\} \notin MB_p^m(f, [x_{j-1}, x_j]). \tag{1.44}$$

Then, according to Lemma 1.31

$$\exists \kappa \in BP(f, (x_{j-1}, x_j)) : v_p(f, [x_{j-1}, x_j]) = s_p(f, \kappa). \tag{1.45}$$

On the other hand, from the Proposition 1.20 we get

$$\begin{aligned} v_p(f, [x_0, x_n]) &= v_p(f, [x_0, x_{j-1}]) + v_p(f, [x_{j-1}, x_j]) + v_p(f, [x_j, x_n]) \\ &= s_p(f, \{x_i\}_{i=0}^{j-1}) + s_p(f, \kappa) + s_p(f, \{x_i\}_{i=j}^{n}). \end{aligned}$$

Thus, $\{x_i\}_{i=0}^{j-1} \cup \kappa \cup \{x_i\}_{i=j}^{n} \in MB_p(f, [a, b])$. Therefore, $\{x_i\}_{i=0}^{n} \notin MB_p^m(f, [a, b])$, because

$$|\{x_i\}_{i=0}^{j-1} \cup \kappa \cup \{x_i\}_{i=j}^{n}| > |\{x_i\}_{i=0}^{n}|, \tag{1.46}$$

This contradicts the initial assumption. ■

1.4 f-join

Definition 1.33 (f-join). Suppose $f : [a, b] \to \mathbb{R}$ is CPM. We will say that points t_a and t_b ($t_a < t_b$) are *f-joined* in interval $[a, b]$ if

$$\exists \{x_j\}_{j=0}^{n} \in MB_p(f, [a, b]) : [t_a, t_b] = [x_{j-1}, x_j], \text{ with some } j. \tag{1.47}$$

Lemma 1.34. Suppose $f : [a, b] \to \mathbb{R}$ is CPM and $a \le t_a < t_b \le b$. If points t_a and t_b are f-joined, then

$$v_p(f, [t_a, t_b]) = |f(t_b) - f(t_a)|^p. \tag{1.48}$$

Proof. Let $f : [a, b] \to \mathbb{R}$ is CMD, $t_a < t_b$ and points t_a and t_b are f-joined. Then exists $\{x_j\}_{j=0}^{n}$ and j from the Definition 1.33. According to definition of the p-variation and properties of s_p function we get

$$v_p(f, [a, b]) = s_p(f, \{x_i\}_{i=0}^{K(f)}) = s_p(f, \{x_i\}_{i=0}^{j-1}) + s_p(f, \{x_{j-1}, x_j\}) + s_p(f, \{x_i\}_{i=j}^{K(f)}). \tag{1.49}$$

Suppose to the contrary that $\exists r \in PP_p(f, [x_{j-1}, x_j]) : s_p(f, r) > s_p(f, \{x_{j-1}, x_j\})$, thus

$$v_p(f, [a, b]) < s_p(f, \{x_i\}_{i=0}^{j-1}) + s_p(f, r) + s_p(f, \{x_i\}_{i=j}^{K(f)}) = s_p(f, \kappa), \tag{1.50}$$

where $\kappa := \{x_i\}_{i=0}^{j-1} \cup r \cup \{x_i\}_{i=j}^{K(f)}$. The last inequality contradicts the definition of p-variation, therefore $\{x_{j-1}, x_j\}$ is the partition of meaningful break points in interval $[x_{j-1}, x_j]$, thus

$$v_p(f, [x_{j-1}, x_j]) = s_p(f, \{x_{j-1}, x_j\}) = |f(x_j) - f(x_{j-1})|^p. \tag{1.51}$$

■

Definition 1.35 (Extremum). We will call the point $t \in [a, b]$ an *extremum* of the function f in interval $[a, b]$ if $f(t) = \sup\{f(z) : z \in [a, b]\}$ or $f(t) = \inf\{f(z) : z \in [a, b]\}$.

Definition 1.36 (Pseudo-monotonic function). A function f is called *pseudo-monotonic* in interval $[a, b]$ if

$$\forall x \in [a, b] : f(x) \in [\min(f(a), f(b)), \max(f(a), f(b))]. \tag{1.52}$$

Proposition 1.37. Let $f : [a, b] \to \mathbb{R}$ be CPM and points t_a and t_b are f-joined. Then points t_a are t_b extremums of the function in the interval $[t_a, t_b]$. In addition, there is no other break point $d \in BP(f, (t_a, t_b))$ that is an extremum of the function in the interval $[t_a, t_b]$.

Proof. Let $f : [a, b] \to \mathbb{R}$ be CPM and points t_a and t_b are f-joined. Since $f \overset{PM}{=} -g$, with out loss of generality we can assume that $f(t_a) \leq f(t_b)$. Suppose to the contrary that $f(t_b)$ is not an extremum of the function in interval $[t_a, t_b]$. Hence, $\exists c \in [t_a, t_b] : f(c) > f(t_b)$. Therefore, $|f(c) - f(t_a)|^p > |f(t_b) - f(t_a)|^p$. According to Proposition 1.34, $v_p(f, [t_a, t_b]) = |f(t_b) - f(t_a)|^p$, thus, $|f(c) - f(t_a)|^p > v_p(f, [t_a, t_b])$, but this contradicts the definition of p-variation. So, point t_b must be an extremum in interval $[t_a, t_b]$. Completely symmetric arguments could be used for point t_a.

In addition, it could be shown that if $d \in BP(f, (t_a, t_b))$, then $f(d) < f(t_b)$. Suppose to the contrary that $\exists d \in BP(f, (t_a, t_b)) : f(d) = f(t_b)$. Then, $|f(d) - f(t_a)|^p = |f(t_b) - f(t_a)|^p$, therefore, $v_p(f, [t_a, d]) = v_p(f, [t_a, t_b])$. Since, $v_p(f, [t_a, t_b]) \geq v_p(f, [t_a, d]) + v_p(f, [d, t_b])$, we get that $v_p(f, [d, t_b]) \leq 0$. So, $v_p(f, [d, t_b]) = 0$, because $v_p(f, [d, t_b]) < 0$ is not valid. According to Lemma 1.3(b), $v_p(f, [d, t_b]) = 0$ iff function is constant in $[d, t_b]$. But this contradicts to the fact that d is a break point (see Lemma 1.24). ■

Corollary 1.38. Let $f : [a, b] \to \mathbb{R}$ be CPM and points t_a and t_b are f-joined. Then function is pseudo-monotonic in interval $[t_a, t_b]$. So, if $f(t_a) \leq f(t_b)$, then

$$\forall x \in (t_a, t_b) : f(t_a) \leq f(x) \leq f(t_b). \tag{1.53}$$

In addition,

$$\forall d \in BP(f, (a, b)) : f(a) < f(d) < f(b). \tag{1.54}$$

Proof. The result follows directly from the Proposition 1.37. ■

Proposition 1.39. Let $f : [a, b] \to \mathbb{R}$ is CPM. If point $d \in [a, b]$ is a break point, which is an extremums of the function f, then d is a meaningful break point.

Proof. Let $f : [a, b] \to \mathbb{R}$ is CPM. Suppose to the contrary that exists $d \in \overline{BP}(f, [a, b])$, which is an extremum of the function, but $d \notin \overline{MB}_p(f, [a, b])$. Let $\{x_j\}_{j=0}^n \in MB(f, [a, b]$ be any partition from $MB(f)$. Since $d \in [a, b]$ and $d \notin \overline{MB}_p(f, [a, b])$,

$$\exists j \in 1, \ldots, n : d \in (x_{j-1}, x_j). \tag{1.55}$$

The points x_{j-1} and x_j are f-joined and d is break point, hence, according Corollary 1.38

$$f(x_{j-1}) < f(d) < f(x_j). \tag{1.56}$$

But this contradicts the assumption that d is the extremum. ■

1.5 Main result

Lemma 1.40. Suppose $C \in \mathbb{R}$, $c_1 \geq 0$ and $1 \leq p < \infty$. Then function $f : [0, \infty) \to \mathbb{R}$ with the values

$$f(x) = (x + c_1)^p - x^p - C, \; x \in [0, \infty), \tag{1.57}$$

are non decreasing in interval $[0, \infty)$.

Proof. Suppose $C \in \mathbb{R}$, $c_1 \geq 0$ and $1 \leq p < \infty$. The the derivative of the function f is

$$\begin{aligned} f'(x) &= p(x + c_1)^{p-1} - px^{p-1} \\ &\geq p(x)^{p-1} - px^{p-1} = 0. \end{aligned}$$

The derivative of function f is non negative, thus the function f is non decreasing. ■

Corollary 1.41. Suppose $c_1 \geq 0$, $1 \leq p < \infty$ and $0 \leq x \leq y$. Then the following implication holds

$$|x + c_1|^p > x^p + C \Rightarrow |y + c_1|^p > y^p + C. \tag{1.58}$$

Proof. Suppose $0 \leq x \leq y$. Since f is non decreasing, $f(x) \leq f(y)$. Therefore, if $f(x) > 0$, then $f(y) > 0$. ■

Proposition 1.42. Let $f : [a, b] \to \mathbb{R}$ is CPM, $1 \leq p < \infty$, $[a', b'] \subset [a, b]$. If

$$\{a', b'\} \in MB_p^m(f, [a', b']), \tag{1.59}$$

then

$$\forall x \in (a', b') : x \notin \overline{MB}_p(f, [a, b]) \tag{1.60}$$

Proof. Suppose to the contrary that the assumptions of the proposition is valid and

$$\exists x \in (a', b') : x \in \overline{MB}_p(f, [a, b]). \tag{1.61}$$

According to the definition of $\overline{MB}_p(f, [a, b])$, the (1.61) means that

$$\exists \{t_i\}_{i=0}^n \in MB_p(f, [a, b]) : x \in \{t_i\}_{i=0}^n. \tag{1.62}$$

Moreover,

$$x \notin \{a, a', b', b\}, \tag{1.63}$$

because $x \in (a', b')$.

Thus, from (1.63) and (1.62) it follows that

$$\exists j \in \{1, 2, \dots, n-1\} : x = t_j. \tag{1.64}$$

Lets denote variables l and k by

$$l := \max \{i \in \{1, 2, \dots, n-1\} : t_i \in (a', b')\}, \tag{1.65}$$
$$k := \min \{i \in \{1, 2, \dots, n-1\} : t_i \in (a', b')\}. \tag{1.66}$$

Since (1.64) holds, the values l and k always exists. According to l and k definitions the following inequality holds

$$t_{k-1} \leq a' < t_k \leq t_l < b' \leq t_{l+1}. \tag{1.67}$$

Points a' and b' are f-joined according to (1.59), therefore, function f is pseudo-monotonic in interval $[a', b']$. Since $v_p(f) = v_p(-f)$ (see Proposition 1.11), with out loss of generality we can assume that interval $[a', b']$ is decreasing, i.e.

$$f(a') > f(b').$$

Moreover, points t_l and t_k are breakpoints, because $\{t_i\}_{i=0}^n \in MB_p(f, [a, b]) \subset BP(f, [a, b])$. Therefore, according to Corollary 1.38 the values $f(t_l)$ and $f(t_k)$ can be only between $f(a')$ and $f(b')$. So,

$$f(a') > f(t_l) > f(b'), \tag{1.68}$$
$$f(a') > f(t_k) > f(b'). \tag{1.69}$$

In addition, since the pairs of points $\{t_l, t_{l+1}\}$ and $\{t_k, t_{k+1}\}$ are also f-joined, the inequalities (1.68) and (1.69) could be extended as follows

$$f(a') > f(t_l) > f(b') \geq f(t_{l+1}), \tag{1.70}$$
$$f(t_{k-1}) \geq f(a') > f(t_k) > f(b'). \tag{1.71}$$

Since $\{a', b'\} \in MB_p^m(f, [a', b'])$, by the Lemma 1.31, any other partition in interval $[a', b']$ is not a meaningful partition, thus

$$|f(a') - f(b')|^p = v_p(f, [a', b']) > v_p(f, [a', t_k]) + v_p(f, [t_k, t_l]) + v_p(f, [t_l, b']). \tag{1.72}$$

According to Lemma 1.3(g), from the following proposition follows

$$\begin{aligned} |f(a') - f(b')|^p &> |f(a') - f(t_k)|^p + v_p(f, [t_k, t_l]) + |f(t_l) - f(b')|^p \\ |f(a') - f(t_l) + f(t_l) - f(b')|^p &> |f(a') - f(t_k)|^p + v_p(f, [t_k, t_l]) + |f(t_l) - f(b')|^p. \end{aligned}$$

From the (1.70) the following inequalities holds

$$\begin{aligned} f(a') - f(t_l) &> 0, \\ f(t_l) - f(t_{l+1}) \geq f(t_l) - f(b') &> 0. \end{aligned}$$

Therefore, we can use Corollary 1.41. According to it, from the inequality (1.73) follows

$$\begin{aligned} |f(a') - f(t_l) + f(t_l) - f(t_{l+1})|^p &> |f(a') - f(t_k)|^p + v_p(f, [t_k, t_l]) + |f(t_l) - f(t_{l+1})|^p \\ |f(a') - f(t_{l+1})|^p &> |f(a') - f(t_k)|^p + v_p(f, [t_k, t_l]) + v_p(f, [t_l, t_{l+1}]). \end{aligned}$$

Absolutely symmetric argument could be used in other direction. Firstly lets modify last inequality

$$|f(a') - f(t_k) + f(t_k) - f(t_{l+1})|^p > |f(a') - f(t_k)|^p + v_p(f, [t_k, t_l]) + v_p(f, [t_l, t_{l+1}]),$$

then form the (1.71) holds the following inequalities

$$\begin{aligned} f(t_k) - f(t_{l+1}) &> 0, \\ f(t_{k-1}) - f(t_k) \geq f(a') - f(t_k) &> 0. \end{aligned}$$

Thus, from Corollary 1.41 we get

$$\begin{aligned} |f(t_{k-1}) - f(t_k) + f(t_k) - f(t_{l+1})|^p &> |f(t_{k-1}) - f(t_k)|^p + v_p(f, [t_k, t_l]) + v_p(f, [t_l, t_{l+1}]) \\ |f(t_{k-1}) - f(t_{l+1})|^p &> v_p(f, [t_{k-1}, t_k]) + v_p(f, [t_k, t_l]) + v_p(f, [t_l, t_{l+1}]) \end{aligned}$$

Finlay, using Lemma 1.28 we conclude that

$$|f(t_{k-1}) - f(t_{l+1})|^p > v_p(f, [t_{k-1}, t_{l+1}]).$$

This contradicts with the definition of p-variation. ■

Proposition 1.43. Suppose $f : [a,b] \to \mathbb{R}$ is CPM, $1 \le p < \infty$ and $x \in [a,b]$. Then the proposition

$$T := \exists (a', b') \in \Big\{ (u,v) \in \mathbb{R}^2 : x \in [u,v] \subset [a,b] \Big\} : x \notin \overline{MB}_p(f, [a', b']) \tag{1.73}$$

is equivalent to

$$S := x \notin \overline{MB}_p(f, [a,b]). \tag{1.74}$$

Proof. Sufficiency. Suppose $x \notin \overline{MB}_p(f, [a,b])$. Then the proposition (1.73) is obviously true then $[a', b'] = [a,b]$.

Necessity. Suppose $f : [a,b] \to \mathbb{R}$ is CPM, $1 \le p < \infty$, $x \in [a,b]$ and (1.73) holds.

Let $\{t_i\}_{i=0}^n \in MB_p^m(f, [a', b'])$. Since $x \in [a', b']$, then $\exists j : x \in (t_{j-1}, t_j)$. In addition, the Lemma 1.32 ensures that $\{t_{j-1}, t_j\} \in MB_p^m(f, [t_{j-1}, t_j])$. Thus, we can apply Proposition 1.42 for interval $[a,b]$ and point x. According to it $x \notin \overline{MB}_p(f, [a,b])$. ■

Corollary 1.44. Since $(T \Leftrightarrow S) \Leftrightarrow (\neg T \Leftrightarrow \neg S)$, we get

$$x \in \overline{MB}_p(f, [a,b]) \Leftrightarrow \forall [a', b'] \subset [a,b], x \in [a', b'] : x \in \overline{MB}_p(f, [a', b']).$$

1.6 Extra conclusions

Proposition 1.45. Suppose $f : [a,b] \to \mathbb{R}$ is CPM, $1 \le p < \infty$. Let $\kappa_{ab} \in MP_p^m(f, [a,b])$, then

$$x \in \overline{MB}_p(f, [a,b]) \Leftrightarrow x \in \kappa_{ab}. \tag{1.75}$$

Proof. Sufficiency. If $x \in \kappa_{ab} \in MP_p^m(f, [a,b])$, then $x \in \overline{MB}_p(f, [a,b])$ according to the definition of $\overline{MB}_p(f, [a,b])$ (see Definition 1.27).

Necessary. Suppose to the contrary that $x \in \overline{MB}(f, [a,b])$, but $x \notin \kappa_{ab} \in MB_p^m(f, [a,b])$. The points of the partition κ_{ab} will be denoted t_i, i.e. $\kappa_{ab} = \{t_i\}_{i=0}^n$. Then

$$\exists j \in \{1, 2, \dots, n\} : x \in (t_{j-1}, t_j).$$

Since $\{(t_{j-1}, t_j\} \in MB_p^m(f, [t_{j-1}, t_j])$, $x \notin \overline{MB}_p(f, [t_{j-1}, t_j])$. Therefore, we can apply Proposition 1.43. From it follows that

$$x \notin \overline{MB}_p(f, [a,b]).$$

This is the contradiction that proves the proposition. ■

Corollary 1.46. According to the definition of MB (see Definition 1.27), from the last proposition directly follows that

$$\forall \kappa \in MB_p(f, [a, b]) : \kappa \subset \kappa_{ab},$$

this means

$$\exists! \kappa_{ab} \in MB_p^m(f, [a, b]).$$

Lemma 1.47. Let $C \geq 0$ and $c_i \geq 0,\ i = 1, \ldots, n$. Suppose $1 \leq p < q < \infty$. Then the implication holds

$$C^p > \sum_{i=1}^{n} c_i^p \Rightarrow C^q > \sum_{i=1}^{n} c_i^q.$$

Proof. Let denote the function $f(p)$ as

$$f(p) = C^p - \left[\sum_{i=1}^{n} c_i^p\right].$$

The derivative of function f is

$$\begin{aligned} f'(p) &= C^p \log p - \left[\sum_{i=1}^{n} c_i^p \log p\right] \\ &= \left(C^p - \left[\sum_{i=1}^{n} c_i^p\right]\right) \log p. \end{aligned}$$

Thus, if $1 \leq p < \infty$ and $C^p > \sum_{i=1}^{n} c_i^p$, then f derivative is non negative. This means that function f is non decreasing, whereas this proof the Lemma 1.47 ■

Proposition 1.48. Suppose $f : [a, b] \to \mathbb{R}$ is CPM and $1 \leq p < q < \infty$. Then

$$\overline{MB}_q(f, [a, b]) \subset \overline{MB}_p(f, [a, b]).$$

Proof. Suppose to the contrary that

$$\exists x \in [a, b] : x \in \overline{MB}_q(f, [a, b]),\ x \notin \overline{MB}_p(f, [a, b]).$$

Let $\{t_i\}_{i=0}^n \in MB_p^m(f, [a, b])$. Because $x \in [a, b]$ and $x \notin \overline{MB}_p(f, [a, b])$, then

$$\exists j \in \{1, 2, \ldots, n\} : x \in (t_{j-1}, t_j). \tag{1.76}$$

Let arbitrary choose $\kappa \in BP(f, (a, b))$. Then according to Lemma 1.31, the following inequality holds

$$|f(t_{j-1}) - f(t_j)|^p > s_p(f, \kappa).$$

Whereas, according to Lemma 1.47,

$$|f(t_{j-1}) - f(t_j)|^q > s_q(f, \kappa).$$

This means that

$$\{t_{j-1}, t_j\} \in MB_q^m(f, [t_{j-1}, t_j]).$$

Finlay, from the Proposition 1.42 follows that $x \notin \overline{MB}_q(f, [a, b])$, but this contradicts the assumptions. ■

2 p-variation calculus

This chapter will present the algorithm that calculates p-variation for the sample (see Def. 1.15). Nonetheless, this algorithm could be used to calculate the p-variation for arbitrary piecewise monotone function. This algorithm will be called *pvar*. It is already realised in the R (see [2]) package *pvar* and is publicly available on CRAN[1].

Suppose $X = \{X_i\}_{i=0}^n$ is any real-value sequence of numbers. We will call such sequence a *sample*, whereas n will be referred to as a *sample size*. The formal definition of p-variation of the sample is given in Definition 1.15, but it is more intuitive to consider a sample as continuous piecewise monotone (CPM) function, namely, Proposition 1.19 states that if $p \geq 1$ then

$$v_p(X) = v_p(L_X(t), [0, n]), \tag{2.1}$$

where $L_X(t)$ is piecewise linear function connecting sample points (Def. 1.16). Since $L_X(t)$ is CPM, all the properties proved in this chapter are applicable to it.

On the other hand, the algorithm does not use $L_X(t)$ function directly, but it actually operates the sample X. Therefore, in the context of algorithm it is more convenient to use the sample X rather then the function $L_X(t)$. Thus in this part we will be using X, but please don't be mislead then we will refer to properties of p-variation that were proved to functions. In that case we actually have in mind function $L_X(t)$.

The $L_X(t)$ is CPM and all its break points could be only at points $t = 0, 1, \ldots, n$, i.e. at points where $L_X(t) = X_0, X_1, \ldots, X_n$. Therefore, according to Prop. 1.26, the p-variation of the sample could be expressed as

$$v_p(X) = \max \left\{ \sum_{i=1}^{k} |X_{j_i} - X_{j_{i-1}}|^p : 0 = j_0 < \cdots < j_k = n,\ k = 1, \ldots, n \right\}. \tag{2.2}$$

[1] http://cran.r-project.org/web/packages/pvar/index.html

It is worth noting, that the procedure *pvar* could be used to calculate the p-variation of any piecewise monotone function. Let assume f is any piecewise monotone function. According to definition (see 1.6) there are points $a = x_1 < \cdots < x_n = b$ for some finite n such that f is monotone on each interval $[x_{j-1}, x_j]$, $j = 1, 2, \ldots, n$. Let the sample X contain the values of the function f at the points $\{x_i\}_{i=0}^n$, namely, $X = \{X_i\}_{i=0}^n := \{f(x_i)\}_{i=0}^n$. It is straight forward to see that $f \overset{PM}{=} L_X$, therefore, if $p \geq 1$ then $v_p(X) = v_p(L_X) = v_p(f)$.

2.1 Main function

The procedure *pvar* that calculates p-variation will be presented here. Firstly, we will introduce the main schema, further, each step will be discussed in more details.

The main principal of the whole procedure is to identify meaningless points using known properties of p-variation. If a single point X_i is identified as meaningless, then we could exude it from further consideration. As long as X_i is meaningless, the actual value of X_i has no effect on p-variation since X_i do not participate in the sum of p-variation. Actually, X_i could be ignored, but it is more convenient to drop it form the sample, because it is hard to apply properties then not all points are under consideration. The *drop* operation means that we updating sample X with new sample X' which do not have X_i element, namely,

$$\{X'_i\}_{i=0}^{n-1} := \{X_0, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n\}.$$

The most important properties which were used in algorithm are ... ??? Later on, we will discus in great detail how we actually using them. The main steps in *pvar* procedure goes as follows.

Procedure *pvar*. Input: sample X, scalar p.

1. *Removing monotonic points.* According to proposition ??? all points that are not the end of monotonic interval could be excluded from further consideration.

2. *Checking all small intervals.* Every small intervals are checked if it is possible to identify meaningless points. All points that are found to be meaningless are excluded. This operation is based on proposition ??? which states that if point is not meaningful in small interval, then it will not be meaningful in big interval as well.

3. *Merging small intervals.*

The corresponding psiaudocode of the main functions goes as follows

```
pvar <- function(x, p){
  partition <- Remove_Monotonic_Points(x)
  partition <- Test_Points_In_Small_Intervals(partition, p, LSI
   = 3)
  partition <- Merge_Small_Intervals(partition, p, LSI=3+1)
  pvar <- sum(abs(diff(partition))^p)
  return(pvar)
}
```

2.2 Detail explanation

2.2.1 Removing monotonic points

According to Proposition 1.12 the p-variation is achieved on the partition of monotonic intervals. Therefore the points that are not the end of monotonic interval could be drop out.

There are multiple options to perform this checking. The method that was actuality applied involves two steps:

1. *Removing constant intervals.* If $X_t = X_{t+1},\ t = 1, \ldots, n-1$ then X_t is removed since both points X_t and X_{t+1} cannot be in break points partition. Let Label the new sample as $\{Y_i\}_{i=0}^{m}$.

2. *Finding break point.* If ΔY_t and ΔY_{t+1} $(t = 1, \ldots, m-1)$ has alternating sign, then point Y_t is considered to be break point. In addition points Y_0 and Y_m is also the break points since end points of the interval is always included in partition (see Definition 1.1). All other point are identified as meaningless and removed from sample.

This procedure can by performed very quickly, since it does not include any interdependent checking and could be done with vectorised functions. In *pvar* package all break points could be found with *ChangePointsId* function.

2.2.2 Test points in small intervals

Todo: papildyti vien savybe su alternuojanciais intervlais

Lemma 2.1. Let $f : [a, b]-> $ be CPM and $\kappa \in X(f)$. If $|\kappa|$ is even then function f has at least one inner meaningful break point.

Proof. Didejantys, pažejantys intervalai. tarki, kad didejantys, tuomet nei a nei b nera maksimumas. Maksimumas yra kažkur ir jis yra reiksmingas. ∎

Lemma 2.2. Turim skaidini, ir ysa istirti visi mazesnieji intervalai iki $n-1$ tada pilnas intervalas yra arba visi arba nei vienas.

Proof. Pritaikom skaidimo teorema. ∎

Proposition 1.43 is very important for speeding up the calculus of p-variation. It states that if the point was not significant in any small interval, then it will not be significant in the full sample as well. Moreover, Proposition 1.20, Lemmas ?? and ?? gives an effective way to identify some meaningless points. Those property allows to drop significant amount of points using quite simple operations.

Suppose X is a sample without monotonic points. Let consider any small interval with $m+1$ points ($m = 2, \dots, n-1$), namely, lets examine a sub-sample $\{X_i\}_{i=j}^{j+m}$ for any $j = 0, ..., n-m$. Based on Proposition 1.20, if

$$\sum_{i=j+1}^{j+m} |X_i - X_{i-1}|^p < |X_j - X_{j+m}|^p, \tag{2.3}$$

then partition $\{j, \dots, j+m\}$ could not be meaningful in interval $[j, j+m]$, therefore, some of the points $\{X_i\}_{i=j+1}^{j+m-1}$ must be meaningless. Those meaningless points can be easily identify if we systemically check all intervals starting from small m.

If $m = 2$ then $|\{X_i\}_{i=j}^{j+2}| = 2$, thus according to Lemma 2.1, $\{X_i\}_{i=j}^{j+m}$ has at least one inner significant point. Since it only has one inner point, then this point must be significant. So, we don't need to do any checkings here.

If $m = 3$ and (2.3) holds, then at least one point must be insignificant. Moreover, according to Lemma 2.2, in this case both middle points X_{t+1} and X_{t+2} must be meaningless.

So, some insignificant points could be identified by applying this kind of checking for all sub-samples $\{X_i\}_{i=j}^{j+3}$, $j = 0, ..., n-3$. Then all identified meaningless points could be droped out. As a result, the whole sample changes, hence, we can apply the same procedure again and try to find even more meaningless points. We should repeat the iterations until no new meaningless points are identified.

If $m = 4$ then $|\{X_i\}_{i=j}^{j+2}|$ is even, thus, according to Lemma 2.1, $\{X_i\}_{i=j}^{j+m}$ has at least one inner significant point. So, all inner points are significant, based on Lemma 2.1. We can conclude, that this argument holds in all cases, then m is even, so we don't need to check it any more.

So, the can go further and apply same checking with $m = 5$. Suppose all sub-samples $\{X_i\}_{i=j}^{j+3}$, $j = 0, ..., n-3$ are already checked. If (2.3) holds, then based on the same Lemma, all middle points X_{t+1}, X_{t+2}, X_{t+3} and X_{t+4} are identified as meaningless and could be dropped out. If we want to apply this checking in new itaration, then we should start checking from $m = 3$ again, because we must make sure that all sub-samples $\{X_i\}_{i=j}^{j+3}$, $j = 0, ..., n-3$ are already checked if we want to apply Lemma ??

We can continue checking by this principle, and actually find p-variation. But it is not effective way, because in any case of drop out we should go back and start checking from $m = 3$, so, it is actually effective for small m. From Monte-Carlo experiment we made a recommendation to check up till $m = 7$.

The final psiaudocode of this procedure is given below

```
SET dum_X to NULL
WHILE dum_X is not equal to X
  SET dum_X = X
  SET d=3
  CHECK all sub-samples with m=d
  UPDATE X by dropping all meaningless points
  WHILE dum_X is equal to X
    SET d = d + 2
    CHECK all sub-samples with m=d points
    UPDATE x by dropping all meaningless points
  ENDWHILE
ENDWHILE
```

2.2.3 Merging small intervals

Adding an extra point to interval Let suppose ... Then x_0 could make a f-join with any other partition point.

Test reasinable posibilities. If find - delete all midpoints

Merge two intervals

```
1. Find potential points from one side an the other

2. Merging
for i in
  AddPoint(i)
EndFor
```

Merging all small intervals

```

while the length of in > 1

  for
    int[i] = merge(int[i], int[i+1])
    i+2;
  end for
  delete all even int



end
```

3 Conclusion

References

[1] J. Qian. The p-variation of Partial Sum Processes and the Empirical Process // Ph.D. thesis, Tufts University, 1997.

[2] R