Strong Results on Weak Derivatives

Bernd Heidergott

Department of Econometrics and Operations Research
Vrije Universiteit Amsterdam

Anniversary Workshop on the occasion of Prof. Georg Pflug's 60th birthday
September 9, 2011

Outline of the Talk

1987, the Beginning

Weak Derivatives in a Nutshell

The Years 1990 to 1996

The Years 1996 to 2011

1987, the Beginning

1987: The "First" Definition

vrije Universiteit amsterdam

254

<u>Definition 1.</u> A function $x \longrightarrow \mu_x$, mapping an open subset of $\mathbb{R}^k$ into $\mathcal{M}_1$ is called weakly differentiable at the point x, if there is a k-vektor of signed finite measures $\mu_x' := (\mu_x'^{(1)}, \quad ,\mu_x'^{(k)})$; $\mu'^{(i)} \in \mathcal{M}$ such that

$$|\langle g,\mu_{x+h}\rangle - \langle g,\mu_x\rangle - \sum h_i \langle g,\mu_x'^{(i)}\rangle| = o(||h||) \qquad (2)$$

for all $g \in C(S)$ as $h \to 0$. Here $o(\cdot)$ may depend on g.

The derivative μ_x' may be represented as

$$\mu_x' = \left[c_1(\dot{\mu}_x^{(1)} - \ddot{\mu}_x^{(1)}),\ c_2(\dot{\mu}_x^{(2)} - \ddot{\mu}_x^{(2)}), \dots, c_k(\dot{\mu}_x^{(k)} - \ddot{\mu}_x^{(k)})\right]$$

where $\dot{\mu}_x^{(i)}, \ddot{\mu}_x^{(i)} \in \mathcal{M}_1$. We do not require that $\dot{\mu}_x^{(i)}$ and $\ddot{\mu}_x^{(i)}$ are orthogonal of each other, bearing however in mind that c_i is minimized if $\dot{\mu}_x^{(i)} \perp \ddot{\mu}_x^{(i)}$. Note that $\langle g,\mu_x'^{(i)}\rangle = 0$ for the constant function $g \equiv 1$, since $\langle g,\mu_x\rangle \equiv 1$.

Source: Lecture Notes in Control and Information Science, IIASA Conference, Sopron Hungary, August 3-7, 1987.

Properties

1) If $x \longrightarrow \mu_x$ and $x \longrightarrow \nu_x$ is differentiable, with derivative $(c, \dot{\mu}_x, \ddot{\mu}_x)$ resp. $(d, \dot{\nu}_x, \ddot{\nu}_x)$, then $x \longrightarrow \alpha\mu_x + (1-\alpha)\nu_x$ is differentiable with derivative

$$\left[\alpha c + (1-\alpha)d,\ \frac{\alpha c\dot{\mu}_x + (1-\alpha)d\dot{\upsilon}_x}{\alpha c + (1-\alpha)d},\ \frac{\alpha c\ddot{\mu}_x + (1-\alpha)d\dot{\upsilon}_x}{\alpha c + (1-\alpha)d}\right]$$

(Note that $\alpha\dot{\mu}_x + (1-\alpha)\dot{\nu}_x$ is in general not orthogonal to $\alpha\ddot{\mu}_x + (1-\alpha)\ddot{\nu}_x$).

2) Under the same assumptions $x \longrightarrow \mu_x * \nu_x$ (convolution) is differentiable with derivative

$$\left(c+d,\ \frac{c}{c+d}\dot{\mu}_x * \nu_x + \frac{d}{c+d}\mu_x * \dot{\nu}_x,\ \frac{c}{c+d}\ddot{\mu}_x * \nu_x + \frac{d}{c+d}\mu_x * \ddot{\nu}_x\right)$$

3) If T is a measurable transformation which maps μ_x onto μ_x^T i.e.

$$\mu_x^T(A) := \mu_x(T^{-1}(A))$$

then $x \longrightarrow \mu_x^T$ is differentiable with derivative $(c, \dot{\mu}_x^T, \ddot{\mu}_x^T)$

Source: Lecture Notes in Control and Information Science, IIASA Conference, Sopron Hungary, August 3-7, 1987.

Weak Derivatives in a Nutshell

The Formal Set-Up

vrije Universiteit amsterdam

Let $(S, \mathcal{S})$ be a measurable space and let $(\mu_\theta : \theta \in \Theta)$ be a family of measures on $(S, \mathcal{S})$, where $\Theta = (a, b) \subset \mathbb{R}$.

Let $\mathcal{D}$ be some set of measurable real-valued mappings defined on S.

Definition

We call μ'_θ the $\mathcal{D}$-derivative of μ_θ if

$$\forall g \in \mathcal{D} : \quad \frac{d}{d\theta} \int g(s)\mu_\theta(ds) = \int g(s)\mu'_\theta(ds).$$

If we take $\mathcal{D} = C_b$, the set of bounded continuous mappings, then the above definition recovers Georg's early definition from 1987.

The Formal Set-Up

Let $(S, \mathcal{S})$ be a measurable space and let $(\mu_\theta : \theta \in \Theta)$ be a family of measures on $(S, \mathcal{S})$, where $\Theta = (a, b) \subset \mathbb{R}$.

Let $\mathcal{D}$ be some set of measurable real-valued mappings defined on S.

Definition

We call μ'_θ the $\mathcal{D}$-derivative of μ_θ if

$$\forall g \in \mathcal{D} : \quad \frac{d}{d\theta} \int g(s) \mu_\theta(ds) = \int g(s) \mu'_\theta(ds).$$

If we take $\mathcal{D} = C_b$, the set of bounded continuous mappings, then the above definition recovers Georg's early definition from 1987.

The Formal Set-Up

Let $(S, \mathcal{S})$ be a measurable space and let $(\mu_\theta : \theta \in \Theta)$ be a family of measures on $(S, \mathcal{S})$, where $\Theta = (a, b) \subset \mathbb{R}$.

Let $\mathcal{D}$ be some set of measurable real-valued mappings defined on S.

Definition

We call μ'_θ the $\mathcal{D}$-derivative of μ_θ if

$$\forall g \in \mathcal{D} : \quad \frac{d}{d\theta} \int g(s) \mu_\theta(ds) = \int g(s) \mu'_\theta(ds).$$

If we take $\mathcal{D} = C_b$, the set of bounded continuous mappings, then the above definition recovers Georg's early definition from 1987.

The Formal Set-Up

Let $(S, \mathcal{S})$ be a measurable space and let $(\mu_\theta : \theta \in \Theta)$ be a family of measures on $(S, \mathcal{S})$, where $\Theta = (a, b) \subset \mathbb{R}$.

Let $\mathcal{D}$ be some set of measurable real-valued mappings defined on S.

Definition

We call μ'_θ the $\mathcal{D}$-derivative of μ_θ if

$$\forall g \in \mathcal{D} : \quad \frac{d}{d\theta} \int g(s)\mu_\theta(ds) = \int g(s)\mu'_\theta(ds).$$

If we take $\mathcal{D} = C_b$, the set of bounded continuous mappings, then the above definition recovers Georg's early definition from 1987.

The Triple Representation

Let μ_θ be a $\mathcal{D}$-differentiable probability measure, with $C_b \subset \mathcal{D}$, then probability measures μ_θ^+ and μ_θ^- and a finite constant c_θ exist such that for all $g \in \mathcal{D}$:

$$\frac{d}{d\theta} \int g(s)\mu_\theta(ds) = c_\theta \left(\int g(s)\mu_\theta^+(ds) - \int g(s)\mu_\theta^-(ds) \right).$$

- Derivatives can be estimated/computed by means of differences of stochastic experiments.
- This concept has been successfully applied for stochastic optimization in operations research.

The Triple Representation

Let μ_θ be a $\mathcal{D}$-differentiable probability measure, with $C_b \subset \mathcal{D}$, then probability measures μ_θ^+ and μ_θ^- and a finite constant c_θ exist such that for all $g \in \mathcal{D}$:

$$\frac{d}{d\theta} \int g(s) \mu_\theta(ds) = c_\theta \left(\int g(s) \mu_\theta^+(ds) - \int g(s) \mu_\theta^-(ds) \right) .$$

- Derivatives can be estimated/computed by means of differences of stochastic experiments.
- This concept has been successfully applied for stochastic optimization in operations research.

The Triple Representation

vrije Universiteit amsterdam

Let μ_θ be a $\mathcal{D}$-differentiable probability measure, with $C_b \subset \mathcal{D}$, then probability measures μ_θ^+ and μ_θ^- and a finite constant c_θ exist such that for all $g \in \mathcal{D}$:

$$\frac{d}{d\theta} \int g(s)\mu_\theta(ds) = c_\theta \left(\int g(s)\mu_\theta^+(ds) - \int g(s)\mu_\theta^-(ds) \right).$$

- Derivatives can be estimated/computed by means of differences of stochastic experiments.
- This concept has been successfully applied for stochastic optimization in operations research.

Example, I

vrije Universiteit amsterdam

First, lets have a look at a probability measure μ_θ that has a nice density.
Let μ_θ be a measure on $\mathbb{R}$ with Lebesgue density f_θ.

If $f_\theta(x)$ is differentiable with respect to θ for $x \in \mathbb{R}$, then

$$\frac{d}{d\theta}\int g(x)\mu_\theta(dx) = \frac{d}{d\theta}\int g(x)f_\theta(x)dx = \int g(x)f'_\theta(x)dx,$$

provided interchanging differentiation and integration is allowed for g.
Now separating the positive and negative part of f'_θ yields

$$\int g(x)f'_\theta(x)dx = \int g(x)\max(f'_\theta(x),0)dx - \int g(x)\max(-f'_\theta(x),0)dx$$

Example, I

First, lets have a look at a probability measure μ_θ that has a nice density.
Let μ_θ be a measure on $\mathbb{R}$ with Lebesgue density f_θ.

If $f_\theta(x)$ is differentiable with respect to θ for $x \in \mathbb{R}$, then

$$\frac{d}{d\theta}\int g(x)\mu_\theta(dx) = \frac{d}{d\theta}\int g(x) f_\theta(x)dx = \int g(x) f'_\theta(x)dx,$$

provided interchanging differentiation and integration is allowed for g.
Now separating the positive and negative part of f'_θ yields

$$\int g(x)f'_\theta(x)dx = \int g(x)\max(f'_\theta(x),0)dx - \int g(x)\max(-f'_\theta(x),0)dx$$

Example, I

First, lets have a look at a probability measure μ_θ that has a nice density.
Let μ_θ be a measure on $\mathbb{R}$ with Lebesgue density f_θ.

If $f_\theta(x)$ is differentiable with respect to θ for $x \in \mathbb{R}$, then

$$\frac{d}{d\theta}\int g(x)\mu_\theta(dx) = \frac{d}{d\theta}\int g(x)f_\theta(x)dx = \int g(x)f'_\theta(x)dx,$$

provided interchanging differentiation and integration is allowed for g.
Now separating the positive and negative part of f'_θ yields

$$\int g(x)f'_\theta(x)dx = \int g(x)\max(f'_\theta(x),0)dx - \int g(x)\max(-f'_\theta(x),0)dx$$

Example, I

First, lets have a look at a probability measure μ_θ that has a nice density.
Let μ_θ be a measure on $\mathbb{R}$ with Lebesgue density f_θ.

If $f_\theta(x)$ is differentiable with respect to θ for $x \in \mathbb{R}$, then

$$\frac{d}{d\theta}\int g(x)\mu_\theta(dx) = \frac{d}{d\theta}\int g(x)f_\theta(x)dx = \int g(x)f'_\theta(x)dx,$$

provided interchanging differentiation and integration is allowed for g.
Now separating the positive and negative part of f'_θ yields

$$\int g(x)f'_\theta(x)dx = \int g(x)\max(f'_\theta(x),0)dx - \int g(x)\max(-f'_\theta(x),0)dx$$

Example, II

vrije Universiteit amsterdam

In a concrete situation easier to use representations can be found.

Let f_θ be the density of the exponential distribution with rate θ and let h_θ be the density of the Erlang-2-distribution with parameter θ, then

$$\int g(x) f'_\theta(x) dx = \frac{1}{\theta}\left(\int g(x) f_\theta(x) dx - \int g(x) h_\theta(x) dx\right)$$

for all $g(x)$ bounded by a polynomial in x (which is a choice for $\mathcal{D}$).

In shorthand notation:

$$\mathrm{Exponential}(\theta)' = \left(\frac{1}{\theta}, \mathrm{Exponential}(\theta), \mathrm{Erlang}(2,\theta)\right).$$

So, one part of the derivative is already given by the nominal experiment.

Example, II

vrije Universiteit amsterdam

In a concrete situation easier to use representations can be found.

Let f_θ be the density of the exponential distribution with rate θ and let h_θ be the density of the Erlang-2-distribution with parameter θ, then

$$\int g(x) f_\theta'(x) dx = \frac{1}{\theta} \left(\int g(x) f_\theta(x) dx - \int g(x) h_\theta(x) dx \right)$$

for all $g(x)$ bounded by a polynomial in x (which is a choice for $\mathcal{D}$).

In shorthand notation:

$$\text{Exponential}(\theta)' = \left(\frac{1}{\theta}, \text{Exponential}(\theta), \text{Erlang}(2, \theta) \right).$$

So, one part of the derivative is already given by the nominal experiment.

Example, II

In a concrete situation easier to use representations can be found.

Let f_θ be the density of the exponential distribution with rate θ and let h_θ be the density of the Erlang-2-distribution with parameter θ, then

$$\int g(x) f'_\theta(x) dx = \frac{1}{\theta} \left(\int g(x) f_\theta(x) dx - \int g(x) h_\theta(x) dx \right)$$

for all $g(x)$ bounded by a polynomial in x (which is a choice for $\mathcal{D}$).

In shorthand notation:

$$\mathrm{Exponential}(\theta)' = \left(\frac{1}{\theta}, \mathrm{Exponential}(\theta), \mathrm{Erlang}(2, \theta) \right).$$

So, one part of the derivative is already given by the nominal experiment.

Example, II

In a concrete situation easier to use representations can be found.

Let f_θ be the density of the exponential distribution with rate θ and let h_θ be the density of the Erlang-2-distribution with parameter θ, then

$$\int g(x) f'_\theta(x) dx = \frac{1}{\theta}\left(\int g(x) f_\theta(x) dx - \int g(x) h_\theta(x) dx\right)$$

for all $g(x)$ bounded by a polynomial in x (which is a choice for $\mathcal{D}$).

In shorthand notation:

$$\mathrm{Exponential}(\theta)' = \left(\frac{1}{\theta}, \mathrm{Exponential}(\theta), \mathrm{Erlang}(2, \theta)\right).$$

So, one part of the derivative is already given by the nominal experiment.

Example, III

Let μ_θ denote the uniform distribution on $[0, \theta]$.

Then it holds for all $g \in C_b$ that

$$\frac{d}{d\theta}\left(\frac{1}{\theta}\int_0^\theta g(x)dx\right) = \frac{1}{\theta}\left(\int_0^\theta g(x)\delta_\theta(dx) - \frac{1}{\theta}\int_0^\theta g(x)dx\right),$$

with $\delta_\theta(\cdot)$ denoting the Dirac measure in θ. In shorthand notation:

$$\mathrm{Uniform}(\theta)' = \left(\frac{1}{\theta}, \mathrm{Dirac}(\theta), \mathrm{Uniform}(\theta)\right).$$

So, again one part of the derivative is already given by the nominal experiment.

Example, III

vrije Universiteit amsterdam

Let μ_θ denote the uniform distribution on $[0, \theta]$.

Then it holds for all $g \in C_b$ that

$$\frac{d}{d\theta}\left(\frac{1}{\theta}\int_0^\theta g(x)dx\right) = \frac{1}{\theta}\left(\int_0^\theta g(x)\delta_\theta(dx) - \frac{1}{\theta}\int_0^\theta g(x)dx\right),$$

with $\delta_\theta(\cdot)$ denoting the Dirac measure in θ. In shorthand notation:

$$\mathrm{Uniform}(\theta)' = \left(\frac{1}{\theta}, \mathrm{Dirac}(\theta), \mathrm{Uniform}(\theta)\right).$$

So, again one part of the derivative is already given by the nominal experiment.

Example, III

Let μ_θ denote the uniform distribution on $[0, \theta]$.

Then it holds for all $g \in C_b$ that

$$\frac{d}{d\theta}\left(\frac{1}{\theta}\int_0^\theta g(x)dx\right) = \frac{1}{\theta}\left(\int_0^\theta g(x)\delta_\theta(dx) - \frac{1}{\theta}\int_0^\theta g(x)dx\right),$$

with $\delta_\theta(\cdot)$ denoting the Dirac measure in θ. In shorthand notation:

$$\text{Uniform}(\theta)' = \left(\frac{1}{\theta}, \text{Dirac}(\theta), \text{Uniform}(\theta)\right).$$

So, again one part of the derivative is already given by the nominal experiment.

Example, III

Let μ_θ denote the uniform distribution on $[0, \theta]$.

Then it holds for all $g \in C_b$ that

$$\frac{d}{d\theta}\left(\frac{1}{\theta}\int_0^\theta g(x)dx\right) = \frac{1}{\theta}\left(\int_0^\theta g(x)\delta_\theta(dx) - \frac{1}{\theta}\int_0^\theta g(x)dx\right),$$

with $\delta_\theta(\cdot)$ denoting the Dirac measure in θ. In shorthand notation:

$$\text{Uniform}(\theta)' = \left(\frac{1}{\theta}, \text{Dirac}(\theta), \text{Uniform}(\theta)\right).$$

So, again one part of the derivative is already given by the nominal experiment.

Questions Raised

vrije Universiteit amsterdam

- What is the added value of weak differentiability?
- What is the best choice of $\mathcal{D}$?

Let's go back to 1987. Did Georg address these questions?

Not really, instead he worked on something bigger: the extension to Markov chains.

Questions Raised

- What is the added value of weak differentiability?
- What is the best choice of $\mathcal{D}$?

Let's go back to 1987. Did Georg address these questions?

Not really, instead he worked on something bigger: the extension to Markov chains.

Questions Raised

- What is the added value of weak differentiability?
- What is the best choice of $\mathcal{D}$?

Let's go back to 1987. Did Georg address these questions?

Not really, instead he worked on something bigger: the extension to Markov chains.

Questions Raised

- What is the added value of weak differentiability?
- What is the best choice of $\mathcal{D}$?

Let's go back to 1987. Did Georg address these questions?

Not really, instead he worked on something bigger: the extension to Markov chains.

Questions Raised

- What is the added value of weak differentiability?
- What is the best choice of $\mathcal{D}$?

Let's go back to 1987. Did Georg address these questions?

Not really, instead he worked on something bigger: the extension to Markov chains.

The Years 1990 to 1996

1990: XV. Symposium on OR

vrije Universiteit amsterdam

Georg gave a lecture at the XV. Symposium on Operations Research (1990, Vienna) on his breakthrough result for Markov chains:

$$\pi'_\theta = \pi_\theta \sum_{n=0}^{\infty} P'_\theta P^n_\theta .$$

Published as: On-Line Optimization of Simulated Markovian Processes Mathematics of Operations Research, 1990, pp. 381-395

This was the first time that I heard of weak derivatives and my interest in the topic brought me to the event on the next slide...

1990: XV. Symposium on OR

Georg gave a lecture at the XV. Symposium on Operations Research (1990, Vienna) on his breakthrough result for Markov chains:

$$\pi'_\theta = \pi_\theta \sum_{n=0}^{\infty} P'_\theta P^n_\theta .$$

Published as: On-Line Optimization of Simulated Markovian Processes Mathematics of Operations Research, 1990, pp. 381-395

This was the first time that I heard of weak derivatives and my interest in the topic brought me to the event on the next slide...

1990: XV. Symposium on OR

Georg gave a lecture at the XV. Symposium on Operations Research (1990, Vienna) on his breakthrough result for Markov chains:

$$\pi'_\theta = \pi_\theta \sum_{n=0}^{\infty} P'_\theta P^n_\theta .$$

Published as: On-Line Optimization of Simulated Markovian Processes
Mathematics of Operations Research, 1990, pp. 381-395

This was the first time that I heard of weak derivatives and my interest in the topic brought me to the event on the next slide...

1992: The Raach Workshop

1992: The Raach Workshop

3.11 Definition. A function $x \mapsto \mu_x$, mapping an open subset of $\mathbb{R}^d$ into P is called *weakly differentiable* at the point x, if there is a d-vector of signed finite measures $\mu'_x := \left(\mu'_{x,1}, \cdots, \mu'_{x,d}\right)$; $\mu'_{x,i} \in \mathcal{M}$ such that

$$\|s\|^{-1} \cdot \left| < H, \mu_{x+s} > - < H, \mu_x > - \sum_{i=1}^{d} s_i \cdot < H, \mu'_{x,i} > \right| \to 0, \qquad (3.8)$$

as $s \to 0$ for all $H \in C_b(R)$, where $s = (s_1, \ldots, s_d) \in \mathbb{R}^d$. As in (??), the derivative μ'_x may be represented as

$$\mu'_x = \left(c_{x,1}\left(\dot{\mu}_{x,1} - \ddot{\mu}_{x,1}\right), c_{x,2}\left(\dot{\mu}_{x,2} - \ddot{\mu}_{x,2}\right), \ldots, c_{x,d}\left(\dot{\mu}_{x,d} - \ddot{\mu}_{x,d}\right)\right) \qquad (3.9)$$

where $\dot{\mu}_{x,i}, \ddot{\mu}_{x,i} \in \mathcal{P}$. We do not require that $\dot{\mu}_{x,i}$ and $\ddot{\mu}_{x,i}$ are orthogonal to each other, bearing however in mind that $c_{x,i}$ is minimized if $\dot{\mu}_{x,i} \perp \ddot{\mu}_{x,i}$. Note that $< \mathbf{1}, \dot{\mu}_{x,i} >= 0$ for the constant function $\mathbf{1}(w) \equiv 1$, since $< \mathbf{1}, \mu_x > \equiv 1$.

We write $\mu'_x = (c_x, \dot{\mu}_x, \ddot{\mu}_x)$ to denote the situation that

$$c_x = \left(c_{x,1}, \ldots, c_{x,d}\right), \dot{\mu}_x = \left(\dot{\mu}_{x,1}, \ldots, \dot{\mu}_{x,d}\right), \ddot{\mu}_x = \left(\ddot{\mu}_{x,1}, \ldots, \ddot{\mu}_{x,d}\right)$$

is the derivative of $x \mapsto \mu_x$ at x in the sense of Definition ??.

Source: Optimization of Simulated Discrete-Event-Processes, Seminar Notes.

1996: The Opus Magnum

1996: The Opus Magnum

Georg Pflug: Optimization of Stochastic Models. The Interface between Simulation and Optimization. Kluwer, Dordrecht, 1996.

The Years 1996 to 2011

A Flourishing Theory

vrije Universiteit amsterdam

- The theory for weak differentiation for Markov chains has been further developed (mainly by others).
- Efficient algorithms, on-line implementations, and even weak Taylor series extensions of Markov chains have been developed.
- A sideline of this research has led to numerical algorithms for approximative computation of Markov chains.

A Flourishing Theory

- The theory for weak differentiation for Markov chains has been further developed (mainly by others).
- Efficient algorithms, on-line implementations, and even weak Taylor series extensions of Markov chains have been developed.
- A sideline of this research has led to numerical algorithms for approximative computation of Markov chains.

A Flourishing Theory

vrije Universiteit amsterdam

- The theory for weak differentiation for Markov chains has been further developed (mainly by others).
- Efficient algorithms, on-line implementations, and even weak Taylor series extensions of Markov chains have been developed.
- A sideline of this research has led to numerical algorithms for approximative computation of Markov chains.

A Flourishing Theory

- The theory for weak differentiation for Markov chains has been further developed (mainly by others).
- Efficient algorithms, on-line implementations, and even weak Taylor series extensions of Markov chains have been developed.
- A sideline of this research has led to numerical algorithms for approximative computation of Markov chains.

...and the Product Rule?

...and the Product Rule?

Here is what the book states on the product rule of weak differentiation:

3.28 Remark. The weak derivative obeys the following rules:

1. Convex combinations. If $x \mapsto \mu_x$ and $x \mapsto \nu_x$ are weakly differentiable with derivative $(c_x, \dot{\mu}_x, \ddot{\mu}_x)$ resp. $(d_x, \dot{\nu}_x, \ddot{\nu}_x)$, then $x \mapsto \alpha\mu_x + (1-\alpha)\nu_x$ is weakly differentiable with derivative

$$\left(\alpha c_x + (1-\alpha) d_x, \frac{\alpha c_x \dot{\mu}_x + (1-\alpha) d_x \dot{\nu}_x}{\alpha c_x + (1-\alpha) d_x}, \frac{\alpha c_x \ddot{\mu}_x + (1-\alpha) d_x \ddot{\nu}_x}{\alpha c_x + (1-\alpha) d_x}\right).$$

(Notice that $\alpha\dot{\mu}_x + (1-\alpha)\dot{\nu}_x$ is in general not orthogonal to $\alpha\ddot{\mu}_x + (1-\alpha)\ddot{\nu}_x$ even if $\dot{\mu}_x \perp \ddot{\mu}_x$ and $\dot{\nu}_x \perp \ddot{\nu}_x$).

2. Convolutions. Under the same assumptions $x \mapsto \mu_x * \nu_x$ (the convolution) is weakly differentiable with derivative

$$(c_x + d_x, \alpha_x \dot{\mu}_x * \nu_x + \beta_x \mu_x * \dot{\nu}_x, \alpha_x \ddot{\mu}_x * \nu_x + \beta_x \mu_x * \ddot{\nu}_x),$$

where $\alpha_x = \frac{c_x}{c_x + d_x}$ and $\beta_x = \frac{d_x}{c_x + d_x}$.

3. Transformations. Let S be a continuous mapping $S : R \to R'$, where R' is some metric space and let μ^S denote the image measure

$$\mu_x^S(A) := \mu_x\left(S^{-1}(A)\right),$$

If $x \mapsto \mu_x$ is weakly differentiable, then $x \mapsto \mu_x^S$ is also weakly differentiable with derivative $(c_x, \dot{\mu}_x^S, \ddot{\mu}_x^S)$.

The Decisive Answer in 2010

vrije Universiteit amsterdam

Let μ_θ and ν_θ be two measures on some measurable space $(S, \mathcal{S})$.

Let $|| \cdot ||_v$ denote the weighted supremum norm with v a mapping that is absolutely integrable with respect to μ_θ and ν_θ, for $\theta \in \Theta$. Product spaces are equipped with the product norm.

Theorem

Let $\mathcal{D}$ be a set of measurable mappings from S on $\mathbb{R}$. If

- μ_θ and ν_θ are $\mathcal{D}$-differentiable,
- $(\mathcal{D}, || \cdot ||_v)$ is a Banach space, then

$$(\mu_\theta \times \nu_\theta)' = \mu_\theta' \times \nu_\theta + \mu_\theta \times \nu_\theta' \, .$$

Remark 1: Let $\mathcal{D}$ be the set of continuous mappings and $v \equiv 1$, then the product rule for C_b-differentiability follows from the above result.

Remark 2: If $(\mathcal{D}, || \cdot ||_v)$ is a Banach space, then $\mathcal{D}$-differentiability of μ_θ implies $|| \cdot ||_v$-Lipschitz continuity of μ_θ.

Source: H. and Leahu, Weak Differentiability of Finite Product Measures, Mathematics of Operations Research, 2010.

The Decisive Answer in 2010

Let μ_θ and ν_θ be two measures on some measurable space $(S, \mathcal{S})$.

Let $|| \cdot ||_v$ denote the weighted supremum norm with v a mapping that is absolutely integrable with respect to μ_θ and ν_θ, for $\theta \in \Theta$. Product spaces are equipped with the product norm.

Theorem

Let $\mathcal{D}$ be a set of measurable mappings from S on $\mathbb{R}$. If

- μ_θ and ν_θ are $\mathcal{D}$-differentiable,
- $(\mathcal{D}, || \cdot ||_v)$ is a Banach space, then

$$(\mu_\theta \times \nu_\theta)' = \mu'_\theta \times \nu_\theta + \mu_\theta \times \nu'_\theta \, .$$

Remark 1: Let $\mathcal{D}$ be the set of continuous mappings and $v \equiv 1$, then the product rule for C_b-differentiability follows from the above result.

Remark 2: If $(\mathcal{D}, || \cdot ||_v)$ is a Banach space, then $\mathcal{D}$-differentiability of μ_θ implies $|| \cdot ||_v$-Lipschitz continuity of μ_θ.

Source: H. and Leahu, Weak Differentiability of Finite Product Measures, Mathematics of Operations Research, 2010.

The Decisive Answer in 2010

vrije Universiteit amsterdam

Let μ_θ and ν_θ be two measures on some measurable space $(S, \mathcal{S})$.

Let $||\cdot||_v$ denote the weighted supremum norm with v a mapping that is absolutely integrable with respect to μ_θ and ν_θ, for $\theta \in \Theta$. Product spaces are equipped with the product norm.

Theorem

Let $\mathcal{D}$ be a set of measurable mappings from S on $\mathbb{R}$. If

- μ_θ and ν_θ are $\mathcal{D}$-differentiable,
- $(\mathcal{D}, ||\cdot||_v)$ is a Banach space, then

$$(\mu_\theta \times \nu_\theta)' = \mu_\theta' \times \nu_\theta + \mu_\theta \times \nu_\theta' \, .$$

Remark 1: Let $\mathcal{D}$ be the set of continuous mappings and $v \equiv 1$, then the product rule for C_b-differentiability follows from the above result.

Remark 2: If $(\mathcal{D}, ||\cdot||_v)$ is a Banach space, then $\mathcal{D}$-differentiability of μ_θ implies $||\cdot||_v$-Lipschitz continuity of μ_θ.

Source: H. and Leahu, Weak Differentiability of Finite Product Measures, Mathematics of Operations Research, 2010.

The Decisive Answer in 2010

Let μ_θ and ν_θ be two measures on some measurable space $(S, \mathcal{S})$.

Let $||\cdot||_v$ denote the weighted supremum norm with v a mapping that is absolutely integrable with respect to μ_θ and ν_θ, for $\theta \in \Theta$. Product spaces are equipped with the product norm.

Theorem

Let $\mathcal{D}$ be a set of measurable mappings from S on $\mathbb{R}$. If

- μ_θ and ν_θ are $\mathcal{D}$-differentiable,
- $(\mathcal{D}, ||\cdot||_v)$ is a Banach space, then

$$(\mu_\theta \times \nu_\theta)' = \mu_\theta' \times \nu_\theta + \mu_\theta \times \nu_\theta' \, .$$

Remark 1: Let $\mathcal{D}$ be the set of continuous mappings and $v \equiv 1$, then the product rule for C_b-differentiability follows from the above result.

Remark 2: If $(\mathcal{D}, ||\cdot||_v)$ is a Banach space, then $\mathcal{D}$-differentiability of μ_θ implies $||\cdot||_v$-Lipschitz continuity of μ_θ.

Source: H. and Leahu, Weak Differentiability of Finite Product Measures, Mathematics of Operations Research, 2010.

The Decisive Answer in 2010

vrije Universiteit amsterdam

Let μ_θ and ν_θ be two measures on some measurable space $(S, \mathcal{S})$.

Let $|| \cdot ||_v$ denote the weighted supremum norm with v a mapping that is absolutely integrable with respect to μ_θ and ν_θ, for $\theta \in \Theta$. Product spaces are equipped with the product norm.

Theorem

Let $\mathcal{D}$ be a set of measurable mappings from S on $\mathbb{R}$. If

- μ_θ and ν_θ are $\mathcal{D}$-differentiable,
- $(\mathcal{D}, || \cdot ||_v)$ is a Banach space, then

$$(\mu_\theta \times \nu_\theta)' = \mu'_\theta \times \nu_\theta + \mu_\theta \times \nu'_\theta \,.$$

Remark 1: Let $\mathcal{D}$ be the set of continuous mappings and $v \equiv 1$, then the product rule for C_b-differentiability follows from the above result.

Remark 2: If $(\mathcal{D}, || \cdot ||_v)$ is a Banach space, then $\mathcal{D}$-differentiability of μ_θ implies $|| \cdot ||_v$-Lipschitz continuity of μ_θ.

Source: H. and Leahu, Weak Differentiability of Finite Product Measures, Mathematics of Operations Research, 2010.

The Future of Weak Derivatives

- Applications to the generator of continuous time Markov chains
- Applications to inhomogeneous Markov chains
- Further development of numerical approximations by means of Taylor polynomials

The Future of Weak Derivatives

vrije Universiteit amsterdam

- Applications to the generator of continuous time Markov chains
- Applications to inhomogeneous Markov chains
- Further development of numerical approximations by means of Taylor polynomials

The Future of Weak Derivatives

- Applications to the generator of continuous time Markov chains
- Applications to inhomogeneous Markov chains
- Further development of numerical approximations by means of Taylor polynomials

The Future of Weak Derivatives

- Applications to the generator of continuous time Markov chains
- Applications to inhomogeneous Markov chains
- Further development of numerical approximations by means of Taylor polynomials

Thank You!

Georg wordt 60 jaar,
lang zal hij leven in de gloria!

Van harte gefeliciteerd!