

Geophysics, Bayesian statistics, and Differential Geometry

**steps towards a more efficient observational data
analysis**

W. Dębski

Inst. Geophys. PAS

Plan referatu

- ◆ Motywacja
- ◆ Problemy odwrotne
- ◆ Podejście Bayesowskie
- ◆ Ku kompletnej teorii

Motywacja

Poznawanie otaczającego świata oparte jest na naszej zdolności obserwacji zjawisk zachodzących wokół nas. Wyniki obserwacji mogą być przedstawiane w różny sposób:

◆ jakościowy:

Patrzcie Państwo, te białe chmurki, jak odmienne!

Zrazu jak stada dzikich gęsi lub łabędzi,

A z tyłu wiatr jak sokół do kupy je pędzi:

....

Wszystkie białe jak srebro, zmieszały się - nagle

Z ich karków rosną maszty, z grzyw szerokie żagle,

Tabun zmienia się w okręt i wspaniale płynie

Cicho, z wolna, po niebios błękitnej równinie! ”

A. Mickiewicz (Pan Tadeusz)

Motywacja

◆ ilościowy:

$$T = 24^{\circ}C$$

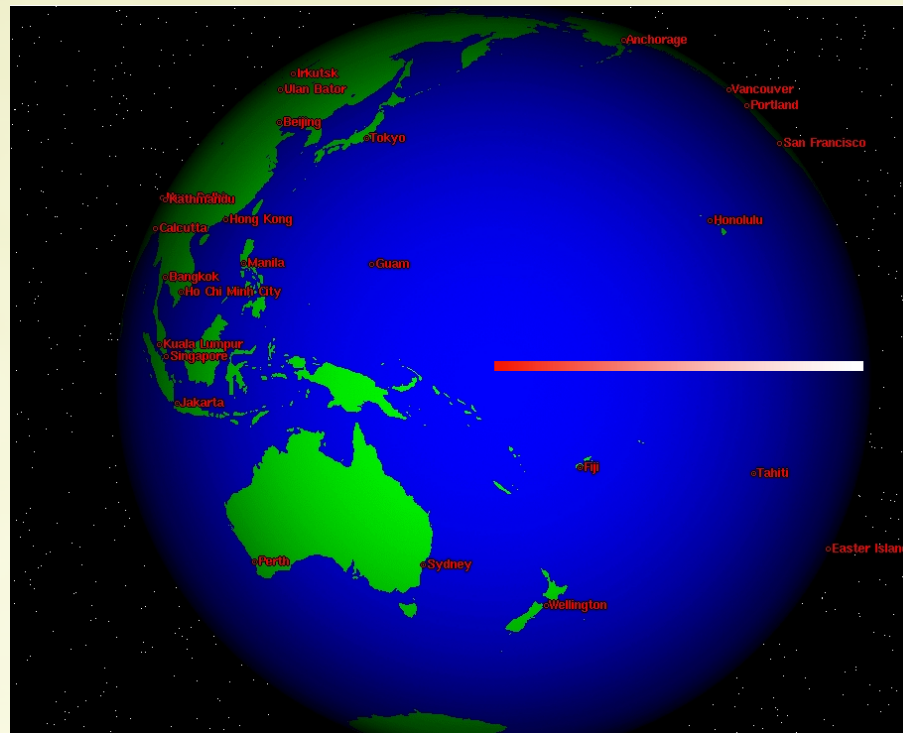
Taki opis rodzi natychmiastowe pytanie:
jaka jest dokładność obserwacji ???

$$T = 24 \pm \Delta T$$

Gauss - nowoczesny rachunek błędów (podstawy)

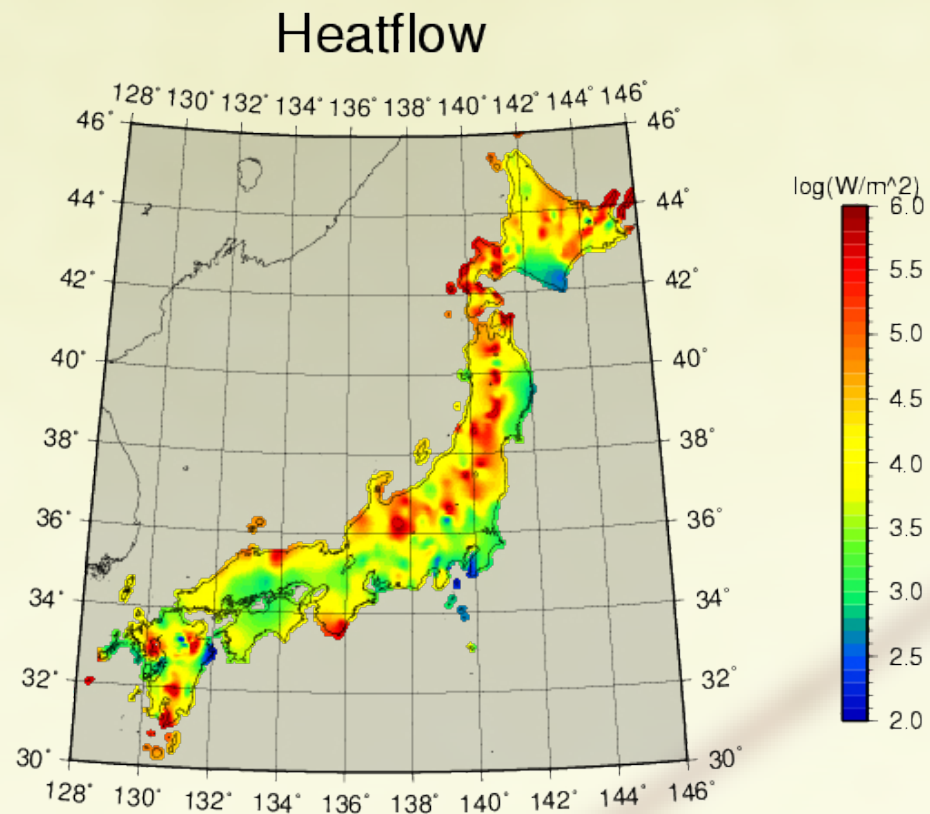
Motywacja

Co zrobić gdy pomiaru wykonać nie możemy ???



Motywacja

Zmierzyć coś co niesie informacje o poszukiwanych wielkościach



Motywacja

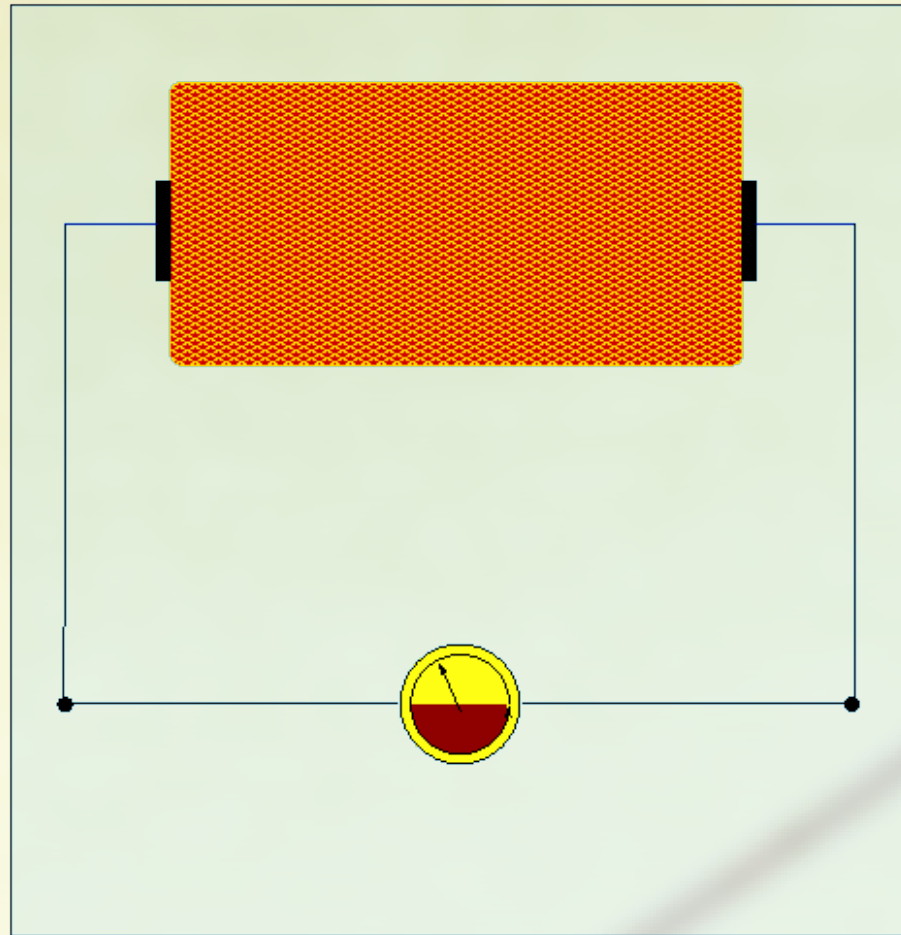
Proces wnioskowania o nieznanymi parametrach na podstawie innych “mierzalnych” wielkości nazywać będą *zagadnieniem odwrotnym*, a zbiór zasad i reguł postępowania *teorią inwersji*.

W takim ujęciu rozwiązanie zagadnienia odwrotnego jest więc *de facto* równoważne wykonaniu pewnego pośredniego pomiaru.

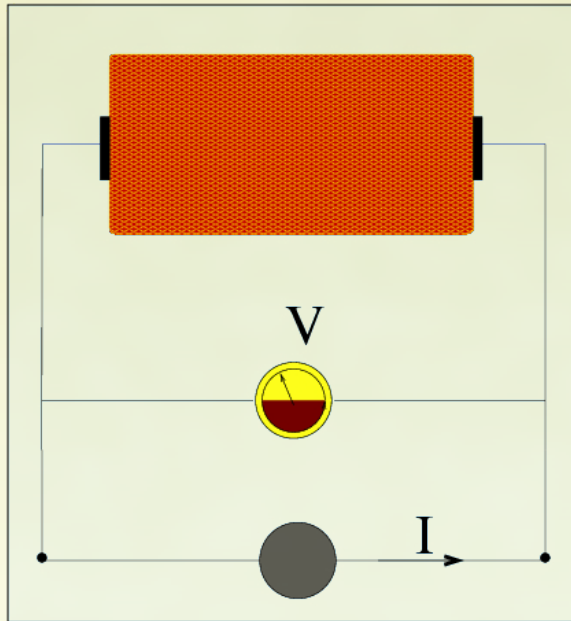
Powraca problem dokładności wyniku.

Przy przedstawionym podejściu jest to jednakże problem znacznie bardziej skomplikowany niż w klasycznej analizie błędów.

Pomiar “prosty” (1)



Pomiar pośredni (1)



Data: V

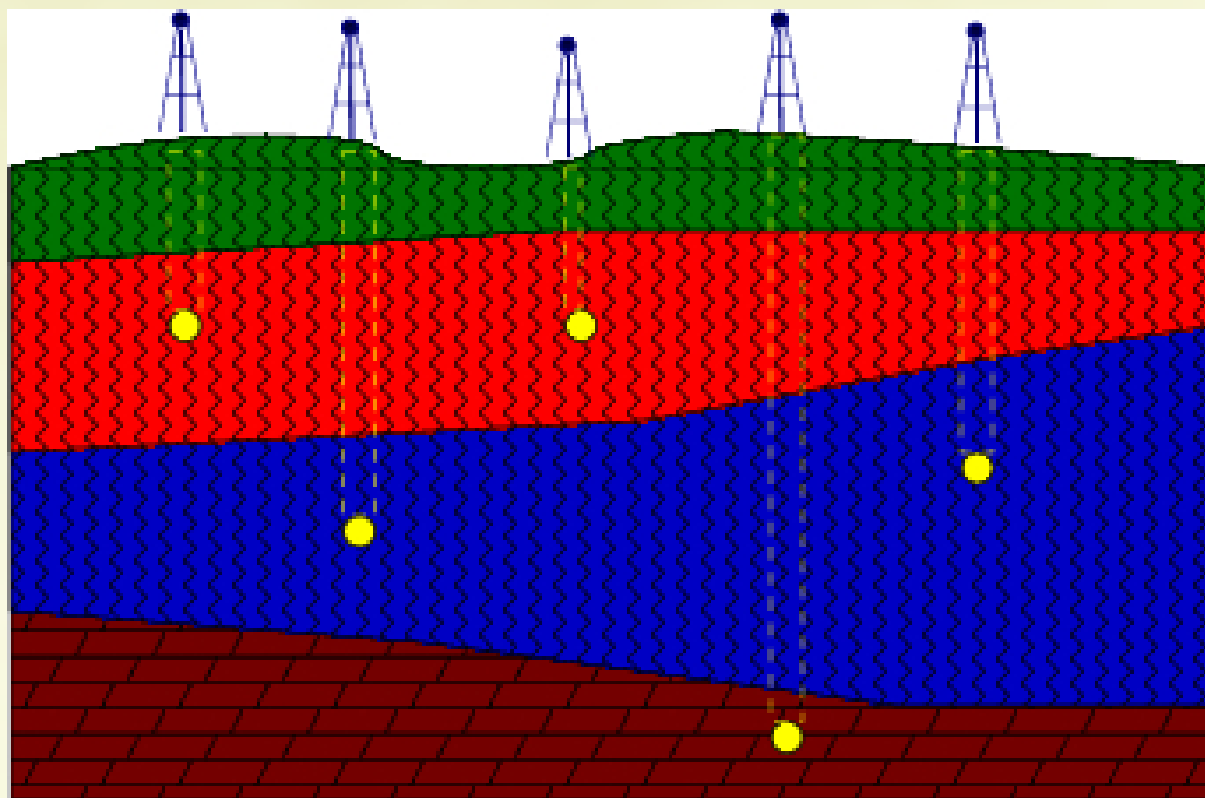
Theory (Ohm law + ...)

$$V = I \rho + n(\rho, \dot{\rho})$$

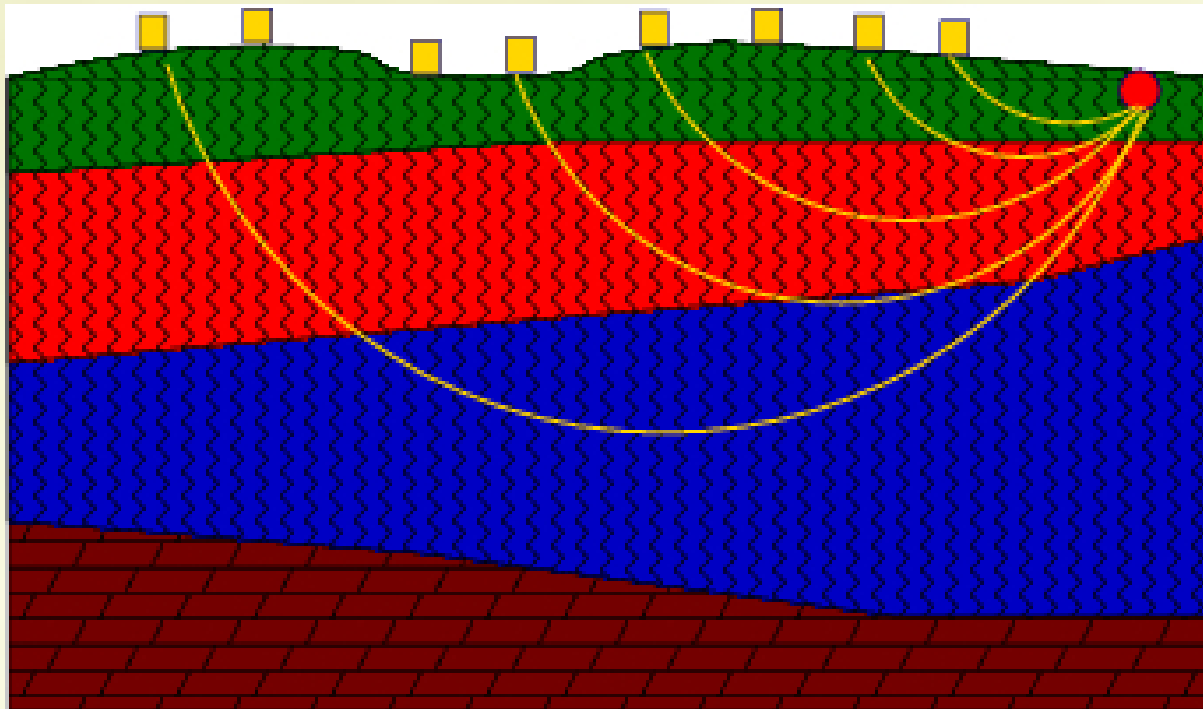
A priori

$$1 < \rho < 10$$

Pomiar józ nie-prosty (2)



Pomiar pośredni (2)

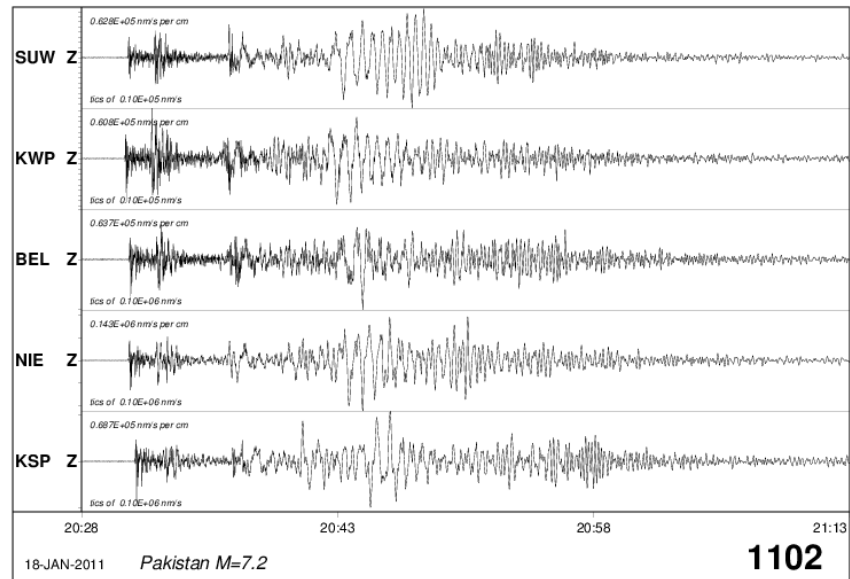


Pomiar pośredni - dane obserwacyjne

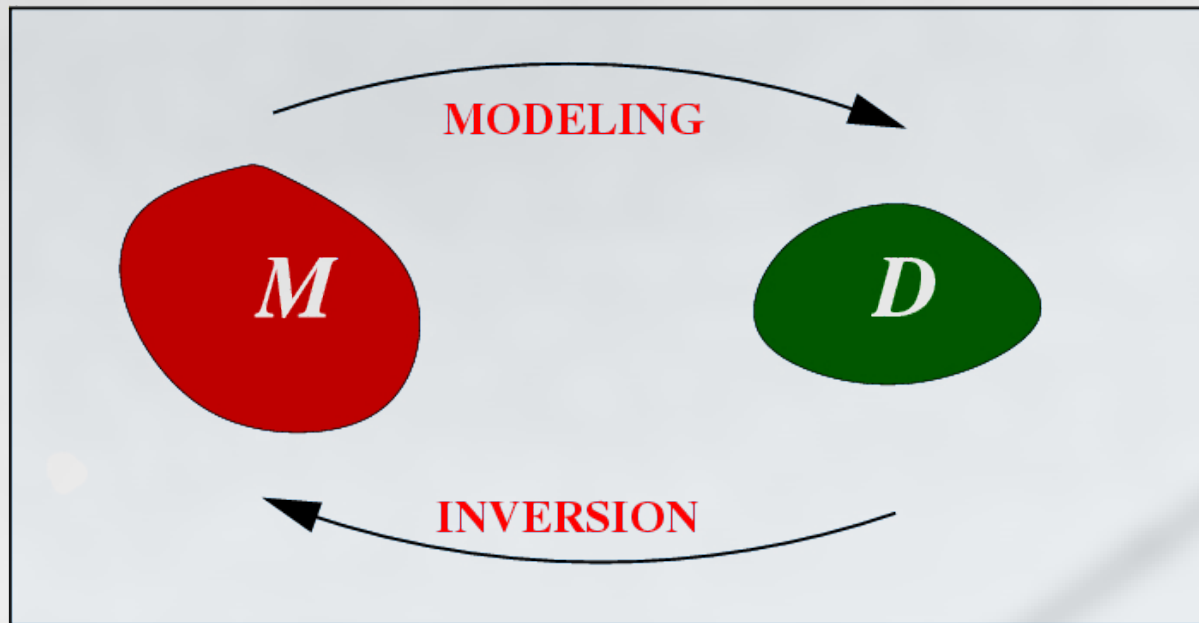
1

$$\left\{ \begin{array}{c} V_1 \\ V_2 \\ \vdots \\ V_N \end{array} \right\}$$

2



Problem odwrotny - ogólnie



Matematyczny opis modelu

System fizyczny:

$$p_1, p_2, \dots, p_K$$

Parametry układu:

$$\mathbf{m} = (m_1, m_2, \dots, m_M)$$

Przewidywane mierzalne wielkości:

$$\mathbf{d} = (d_1, d_2, \dots, d_N)$$

Parametry “ustalone” (znane *a priori*):

$$\mathbf{m}^{ust} = (u_1, u_2, \dots)$$

Problem odwrotny - “szacowanie” parametrów

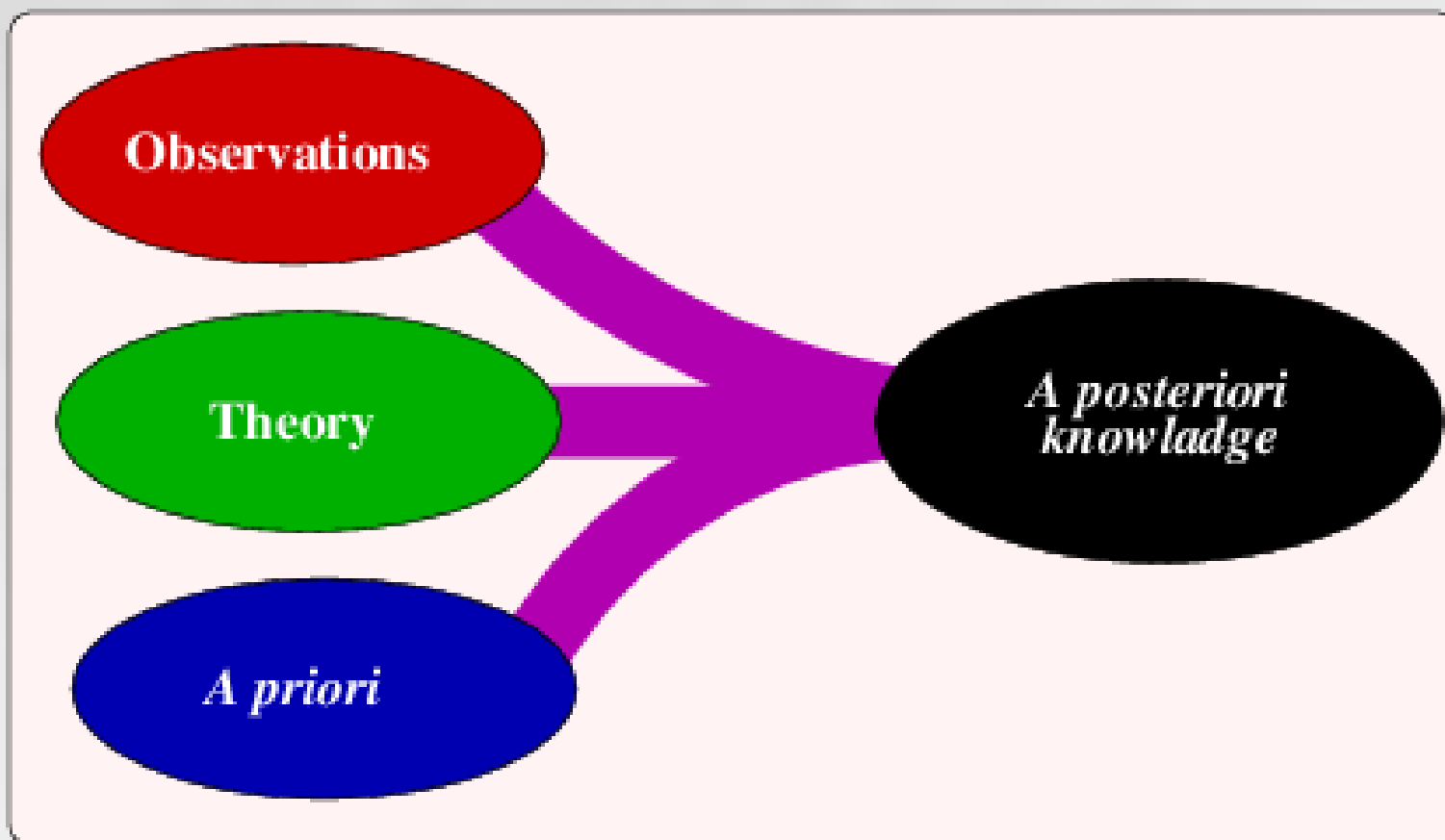
Modelowanie:

$$\mathbf{d}^{th} = f(\mathbf{m}, \mathbf{m}^{ust})$$

Inwersja :

$$\mathbf{d}^{obs} \Rightarrow \mathbf{m} = \dots \pm \dots$$

Problem odwrotny - uogólniamy dalej ...



Liniowy problem odwrotny

$$\mathbf{m} = (m_1, m_2, \dots, m_M) \in \mathcal{M}$$

$$\mathbf{d} = (d_1, d_2, \dots, d_N) \in \mathcal{D}$$

$$\mathbf{d}(\mathbf{m}) = \mathbf{G} \cdot \mathbf{m}$$

$$\mathbf{d}(\mathbf{m}) = \mathbf{G}(\mathbf{m}) \approx \mathbf{G}(\mathbf{m}_o) + \mathbf{H} \cdot \delta\mathbf{m}$$

$$\mathbf{m} = \mathbf{m}_o + \delta\mathbf{m}$$

Liniowy problem odwrotny

$$\mathbf{m} \sim \mathbf{m}^{apr}; \quad \bar{\mathbf{d}} = \mathbf{G} \cdot \mathbf{m}^{apr}$$

$$\mathbf{d} = \mathbf{G} \cdot \mathbf{m} \quad / \quad - \bar{\mathbf{d}}$$

$$\mathbf{d} - \bar{\mathbf{d}} = \mathbf{G} \cdot (\mathbf{m} - \mathbf{m}^{apr})$$

$$\mathbf{G}^T \cdot (\mathbf{d} - \bar{\mathbf{d}}) = \mathbf{G}^T \mathbf{G} \cdot (\mathbf{m} - \mathbf{m}^{apr})$$

$$\mathbf{G}^T \mathbf{G} \implies \mathbf{G}^T \mathbf{G} + \lambda \mathbf{I}$$

$$\mathbf{m}^{est}(\lambda) = \mathbf{m}^{apr} + (\mathbf{G}^T \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^T \cdot (\mathbf{d}^{obs} - \mathbf{G} \cdot \mathbf{m}^{apr})$$

Liniowy problem odwrotny - ocena rozdzielczości

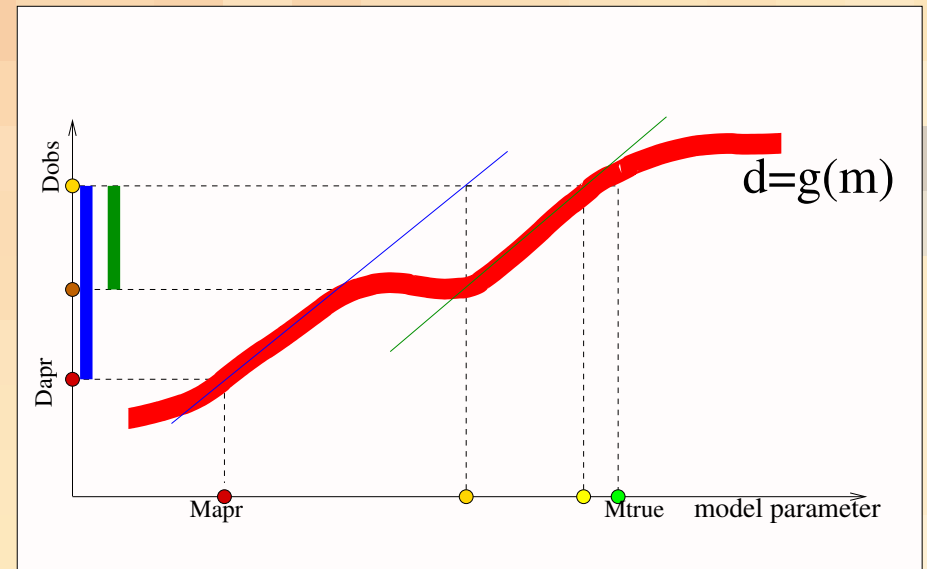
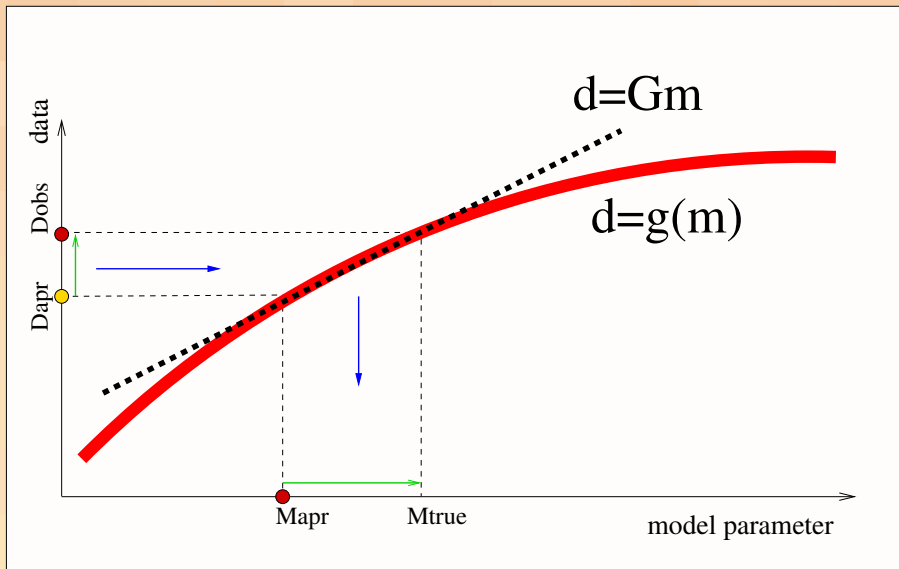
$$\mathbf{d}^{true} = \mathbf{G} \cdot \mathbf{m}^{true}; \quad \mathbf{d}^{obs} = \mathbf{d}^{true}$$

$$\mathbf{m}^{est} - \mathbf{m}^{apr} = \mathbf{R}(\lambda) \cdot (\mathbf{m}^{true} - \mathbf{m}^{apr})$$

$$\mathbf{R} = \frac{\mathbf{G}^T \mathbf{G}}{\mathbf{G}^T \mathbf{G} + \lambda \mathbf{I}}$$

Otrzymane parametry \mathbf{m}^{est} są “przefiltrowanymi” obrazami nieznanymi, prawdziwych \mathbf{m}^{true} . Obecność tego “filtru” wynika ze skończonej ilości posiadanych danych a więc i informacji o parametrach \mathbf{m} .

Problemy odwrotne nieliniowe



Inwersja - metoda optymalizacyjna

$$\mathbf{d}^{th}(\mathbf{m}) = g(\mathbf{m})$$

Optymalizacja: szukamy \mathbf{m}^{est} takiego, że

$$\|\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{est})\| = \min$$

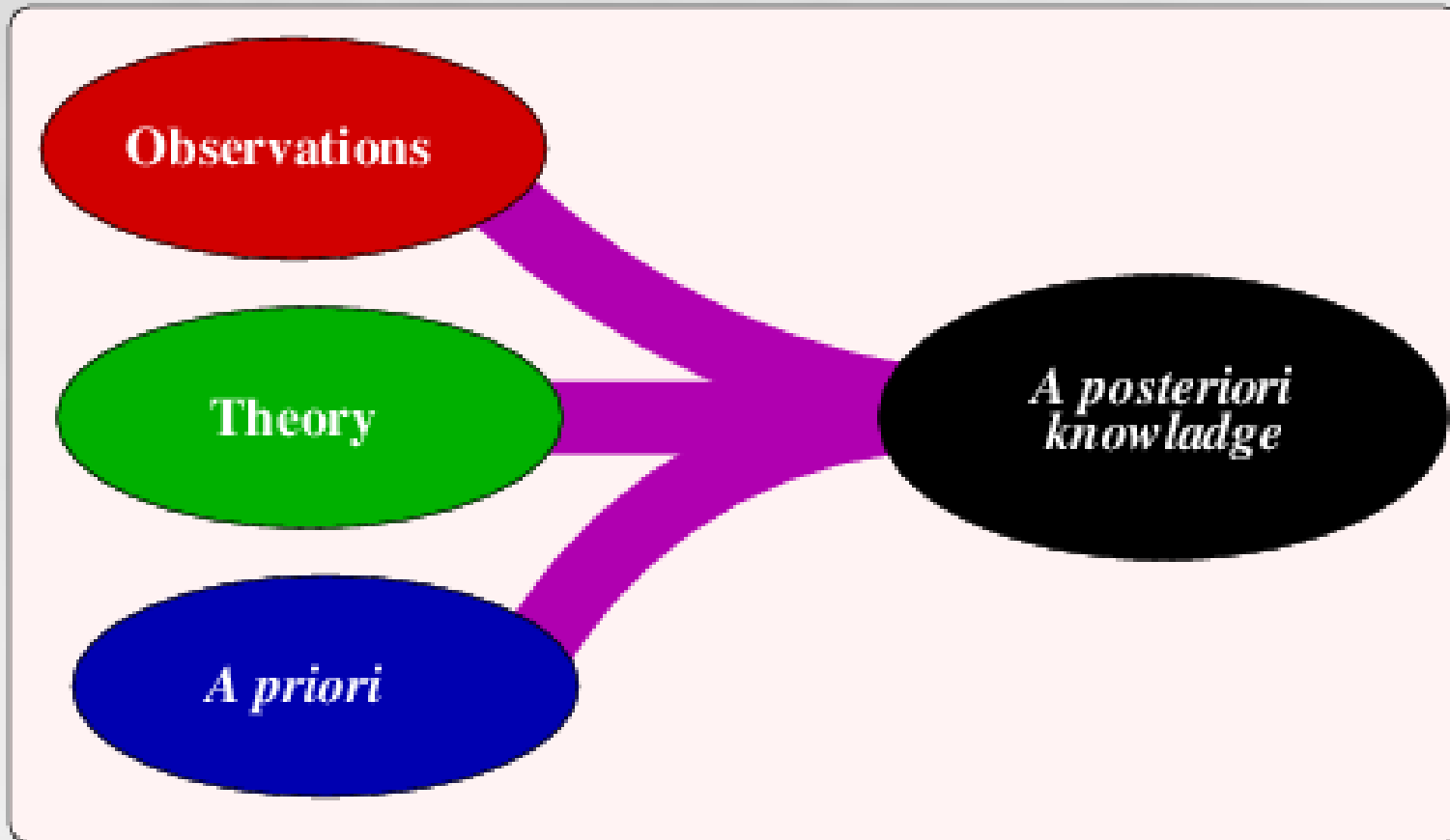
lub

$$\|\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{est})\|_D + \lambda \|\mathbf{m}^{est} - \mathbf{m}^{apr}\|_M = \min$$

Inwersja - metoda optymalizacyjna

- ◆ metoda nieliniowa
- ◆ fizyczny sens regularyzacji
- ◆ poszukujemy jednego (optymalnego) rozwiązania
- ◆ zaniebane wszystkie błędy i niedokładności
- ◆ jak mierzyć dokładność dopasowania $\mathbf{d}^{th} - \mathbf{d}^{obs}$?
- ◆ czy rozwiązanie jest jednoznaczne ?

Wnioskowanie Bayesowskie



Przykłady - uogólnienie

pytanie



dane obserwacyjne

relacje teoretyczne



wnioskowanie



wynik

Probabilistyczny opis informacji

posiadana
informacja \Leftrightarrow rozkład
prawdopod.

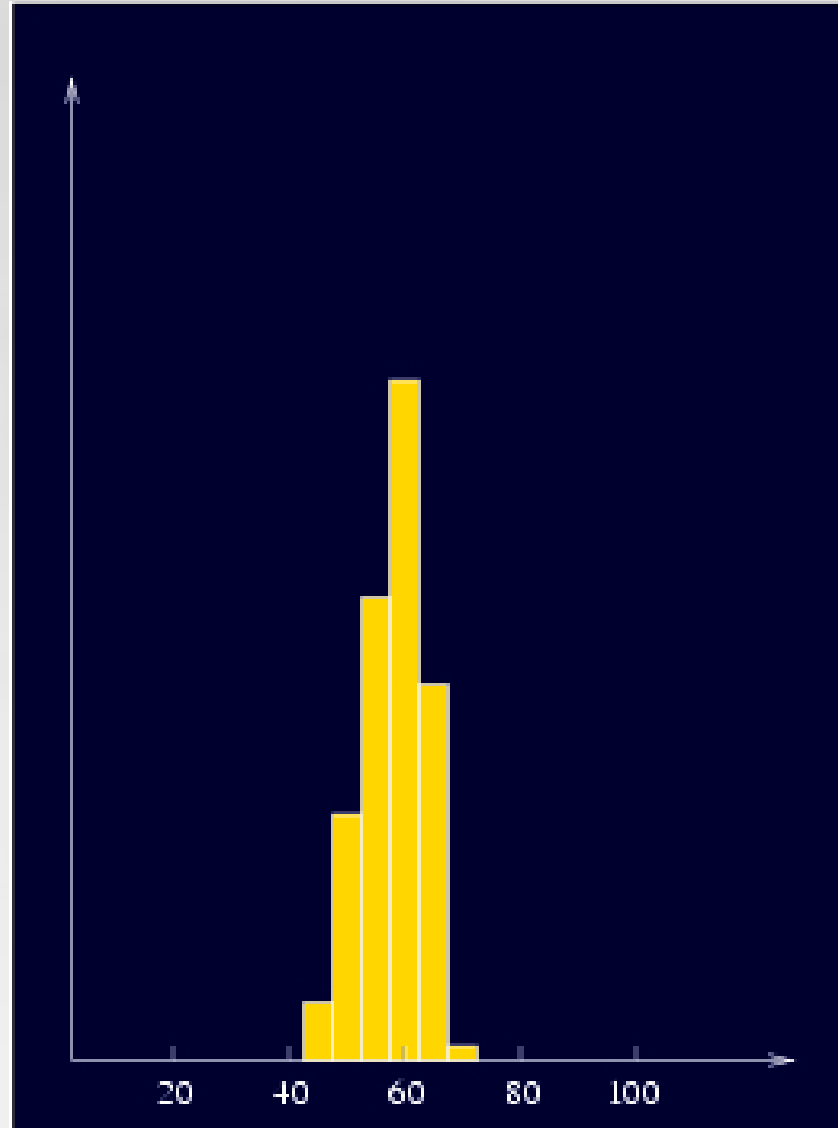
Postulat (Bayes):

Każdą posiadaną informację możemy ilościowo opisać przy pomocy pewnego rozkładu prawdopodobieństwa:

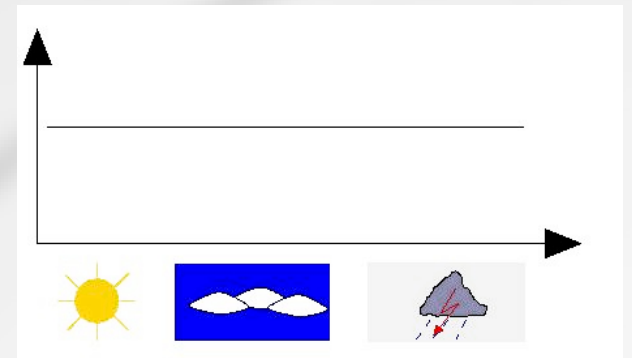
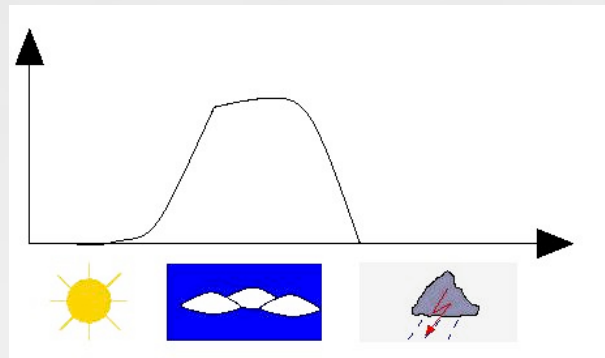
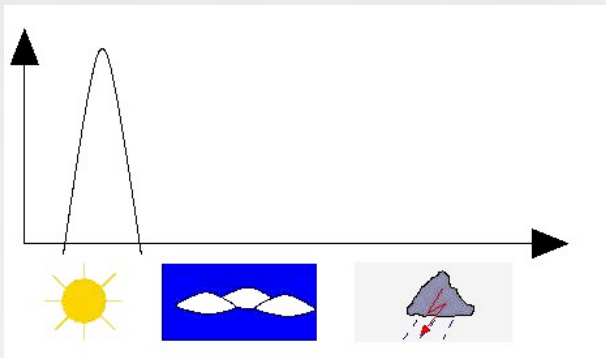
Przykład (1) (Wiek A. Tarantoli ???)



Przykład (1) rozwiązanie



Przykład (2) pogoda



Skoro posiadane informacje opisujemy rozkładami prawdopodobieństwa to jak skonstruować rozkład *a posteriori*?

Twierdzenie Bayesa

$$p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d}) = \frac{p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}|\mathbf{m})p_{\mathbf{m}}(\mathbf{m})}{p_{\mathbf{d}}(\mathbf{d})}$$

$$p_{\mathbf{m}|\mathbf{d}}(\mathbf{m}|\mathbf{d}^{obs}) = \frac{p_{\mathbf{d}|\mathbf{m}}(\mathbf{d}^{obs}|\mathbf{m})p_{\mathbf{m}}(\mathbf{m})}{p_{\mathbf{d}}(\mathbf{d}^{obs})}$$

$$p_{post}(\mathbf{m}|\mathbf{d}^{obs}) \sim p(\mathbf{d}^{obs}|\mathbf{m}) p_{apr}(\mathbf{m})$$

Podejście Tarantoli

Dwa rozkłady opisujące dwie różne informacje

$$p(x), q(x), \quad p, q \in \mathcal{P}$$

$$\mathbf{p} \wedge \mathbf{q}(\mathbf{x}) = \frac{\mathbf{p}(\mathbf{x}) \mathbf{q}(\mathbf{x})}{\mu(\mathbf{x})}$$

Opisuje łączne posiadane informacje

Rozwiązanie problemu inwersyjnego

1. obserwacje: $p(\mathbf{d})$
 2. teoria : $q(\mathbf{m}, \mathbf{d})$
 3. *a priori* : $f(\mathbf{m}, \mathbf{d})$
-

$$(p \wedge q)(\cdot) = \frac{p(\cdot) q(\cdot)}{\mu(\cdot)} \quad \mathcal{P} = \mathcal{M} \times \mathcal{D}$$

$$p(\mathbf{d}) \implies \bar{p}(\mathbf{m}, \mathbf{d}) = \mu(\mathbf{m})p(\mathbf{d})$$

Rozkład *A posteriori* pdf na $\mathcal{M} \times \mathcal{D}$

Trzy kroki:

1. Uogólnienie wszystkich rozkładów na przestrzeń $\mathcal{M} \times \mathcal{D}$
2. Utworzenie uogólnionej funkcji podobieństwa:

$$\mathcal{L}(\mathbf{m}, \mathbf{d}) = (p \wedge q)(\mathbf{m}, \mathbf{d}) = \frac{p(\mathbf{m}, \mathbf{d}) q(\mathbf{m}, \mathbf{d})}{\mu(\mathbf{m}, \mathbf{d})}$$

3. utworzenie rozkładu *A posteriori*:

$$\sigma(m, d) \approx (\mathcal{L} \wedge f)(\mathbf{m}, \mathbf{d}) = \frac{\mathcal{L}(\mathbf{m}, \mathbf{d}) f(\mathbf{m}, \mathbf{d})}{\mu(\mathbf{m}, \mathbf{d})}$$

A posteriori pdf

$$\sigma(\mathbf{m}, \mathbf{d}) = \frac{p(\mathbf{m}, \mathbf{d}) q(\mathbf{m}, \mathbf{d}) f(\mathbf{m}, \mathbf{d})}{\mu^2(\mathbf{m}, \mathbf{d})}$$

Brzegowe rozkłady *A posteriori*

$$\sigma_m(m) = \int_D \sigma(m, d) dD$$

$$\sigma_d(d) = \int_M \sigma(m, d) dM$$

A posteriori pdf

$$\sigma_m(m) = f(m) \cdot L(m, \mathbf{d}^{obs})$$

$$L(\mathbf{m}, \mathbf{d}^{obs}) = \int_D p(\mathbf{d}, \mathbf{d}^{obs}) \frac{q(\mathbf{m}, \mathbf{d})}{\mu(\mathbf{m}, \mathbf{d})} d\mathbf{d}$$

Przykład - dokładna teoria

$$q(\mathbf{m}, \mathbf{d}) = \delta(\mathbf{d} - \mathbf{G}(\mathbf{m})) \mu_M(\mathbf{m})$$

$$p(\mathbf{d}, \mathbf{d}^{obs}) = \rho(\|\mathbf{d} - \mathbf{d}^{obs}\|)$$

wtedy

$$\sigma(\mathbf{m}) = f(\mathbf{m}) \frac{p(\mathbf{d}^{obs} - \mathbf{G}(\mathbf{m}))}{\mu_D(\mathbf{d})}$$

funkcja podobieństwa opisuje niepewność wyznaczenia \mathbf{m} z względu na obecność błędów pomiarowych

Przykład - dokładne pomiary

$$p(\mathbf{d}, \mathbf{d}^{obs}) = \delta(\mathbf{d} - \mathbf{d}^{obs})$$

wtedy

$$\sigma(\mathbf{m}) = f(\mathbf{m}) \frac{q(\mathbf{m}, \mathbf{d}^{obs})}{\mu(\mathbf{m}, \mathbf{d}^{obs})}$$

funkcja podobieństwa opisuje niepewność wyznaczenia \mathbf{m} z względu na obecność błędów modelowania

Przykład - brak informacji teoretycznych

rola informacji *a priori*

$$q(\mathbf{m}, \mathbf{d}) = \mu(\mathbf{m}, \mathbf{d})$$

$$L(\mathbf{m}) = \int_D p(\mathbf{d}, \mathbf{d}^{obs}) \frac{q(\mathbf{m}, \mathbf{d})}{\mu(\mathbf{m}, \mathbf{d})} d\mathbf{d} = \int_D p(\mathbf{d}) d\mathbf{d} = \text{const.}$$

$$\sigma(\mathbf{m}) = f(\mathbf{m})$$

Przykład - brak danych pomiarowych

rola informacji *a priori*

$$p(\mathbf{d}) = \mu_D(\mathbf{d})$$

wtedy

$$L(\mathbf{m}) = \int_D p(\mathbf{d}) / \mu_D(\mathbf{d}) \frac{q(\mathbf{m}, \mathbf{d})}{\mu_M(\mathbf{m})} d\mathbf{d}$$

gdy

$$q(\mathbf{m}, \mathbf{d}) = \mu_M(\mathbf{m}) \rho(\mathbf{d} - G(\mathbf{m}))$$

$$\sigma(\mathbf{m}) = f(\mathbf{m})$$

Problemy

◆ Rozkład nie-informacyjny $\mu(\mathbf{m}, \mathbf{d})$ - to rozkład który opisuje sytuację w której nic nie wiemy. Jak go skonstruować?.

◆ konstrukcja $\mu(\mathbf{m}, \mathbf{d})$ dla innych parametrów (np tylko dodatnich) ???

◆ Stan “pewności”:

$$p_e(\mathbf{m}) = \delta(\mathbf{m} - \mathbf{m}_a)$$

◆ Stan ‘niemożliwy’ :

$$p_n = 0$$

◆ Problem z operacją łączenia informacji

$$p \wedge q(x) = \frac{p(x) q(x)}{\mu(x)}$$

$$p \wedge p(x) = \frac{p^2(x)}{\mu(x)} \neq p(x)$$

co jest sprzeczne ze zdrowym rozsądkiem i klasyczną logiką

- ◆ Inne problemy: $p \wedge q(x)$ nie zachowuje normalizacji i wyprowadza poza przestrzeń rozkładów prawdopodobieństw - może się zdarzyć $p \wedge q = 0$ - nienormalizowalne.
- ◆ *A gdzie rozkłady *apriori* odnośnie teorii???*
- ◆ Niezmienniczość i symetria stanu null-information ($\mu(x)$)

Problemy - brak ogólności

$$\sigma(\mathbf{m}, \mathbf{d}) = \frac{p(\mathbf{m}, \mathbf{d}) q(\mathbf{m}, \mathbf{d}) f(\mathbf{m}, \mathbf{d})}{\mu^2(\mathbf{m}, \mathbf{d})}$$

$$\mathbf{d} = \mathbf{G}(\mathbf{m}) + \epsilon(\mathbf{m}) \Rightarrow q(\mathbf{m}, \mathbf{d}) = q_\epsilon(\epsilon)$$

Gdyby rozważać problem typu “model selection” czyli estymacja $\mathbf{G}(\cdot)$

$$\mathcal{M} \times \mathcal{D} \Rightarrow \mathcal{M} \times \mathcal{D} \times \mathcal{G}$$

Rozszerzenie pomysłu ...

Niech \mathcal{J} będzie przestrzenią parametrów: $\mathcal{J} = \mathcal{M} \times \mathcal{D}$
a $\tilde{\mathcal{P}}$ przestrzenią wszystkich rozkładów prawdopodobieństwa na \mathcal{J}

$$\tilde{\mathcal{P}}_i : \mathcal{J} \rightarrow R^+$$

Rozszerzmy następnie przestrzeń $\tilde{\mathcal{P}}$ jak następuje:

$$\tilde{\mathcal{P}} \rightarrow \tilde{\mathcal{P}} + \{\nu(\cdot)\} + \{\delta(\cdot)\}$$

gdzie $\nu(\mathbf{x}) \equiv 0$

a $\delta(\mathbf{x})$ - jest delta Diraca

Rozszerzenie pomysłu ...

Następnie wprowadźmy klasy abstrakcji (“uogólnione prawdopodobieństwa”)

$$\tilde{\mathcal{P}} \rightarrow \mathcal{P} = \tilde{\mathcal{P}} / R^+$$

co oznacza utożsamienie prawdopodobieństwo “o tym samym kształcie”

$$\exists_{a, a \neq 0} \forall_x \sigma_j(\mathbf{x}) = a\sigma_i(\mathbf{x}) \implies \sigma_j \equiv \sigma_i$$

Rozszerzenie pomysłu ...

Zdefiniujmy operację binarną $\wedge : \mathcal{P} \times \mathcal{P} \longrightarrow \mathcal{P}$

$$(\sigma_i \wedge \sigma_j)(\mathcal{A}) = \begin{cases} \frac{\sigma_i(\mathcal{A})\sigma_j(\mathcal{A})}{\mu(\mathcal{A})} & \text{if } \sigma_i \neq \sigma_j \\ \sigma_i(\mathcal{A}) & \text{if } \sigma_i = \sigma_j \end{cases}$$

$$(\sigma_i \wedge \sigma_j)(\mathbf{x}) = \begin{cases} \frac{\sigma_i(\mathbf{x})\sigma_j(\mathbf{x})}{\mu(\mathbf{x})} & \text{if } \sigma_i \neq \sigma_j \\ \sigma_i(\mathbf{x}) & \text{if } \sigma_i = \sigma_j \end{cases}$$

gdzie

$\mathbf{x} \in \mathcal{J}$, a $\mu(\mathbf{x})$ jest rozkładem nie-informatycznym na \mathcal{J} space.

Rozszerzenie pomysłu ...

Operacja \wedge ma następujące własności:

- a) $(\sigma_i \wedge \sigma_i)(\mathbf{x}) = \sigma_i(\mathbf{x})$ idempotency
- a) $(\sigma_i \wedge \sigma_j) \wedge \sigma_k = \sigma_i \wedge (\sigma_j \wedge \sigma_k)$ associativity
- b) $\sigma_i \wedge \mu = \mu \wedge \sigma_i = \sigma_i$ identity element
- c) $\sigma_i \wedge \sigma_j = \sigma_j \wedge \sigma_i$ commutativity.

Jednym słowem struktura (\mathcal{P}, \wedge) jest półgrupą abelową

Rozszerzenie pomysłu ...

Ponadto

$$\blacklozenge (\sigma_i \wedge \sigma_j)(\mathcal{A}) \neq 0 \implies \sigma_i(\mathcal{A}) \neq 0 \quad \text{and} \quad \sigma_j(\mathcal{A}) \neq 0$$

$$\blacklozenge (\sigma_i \wedge \sigma_j)(\mathbf{y}) = \left| \frac{\partial(x^1, x^2, \dots)}{\partial(y^1, y^2, \dots)} \right| (\sigma_i \wedge \sigma_j)(\mathbf{x})$$

gdzie $|\cdot|$ jest wyznacznikiem macierzy Jacobiego

$$\blacklozenge (\sigma_i \wedge \sigma_{e_i})(\cdot) = \sigma_{e_i}(\cdot)$$

$$\blacklozenge (\sigma_i \wedge \nu)(\cdot) = \nu(\cdot)$$

Rozszerzenie pomysłu ...

Spełnione (?) są też warunki Kołomogorowa

$$\text{a) } (\sigma_i \wedge \sigma_j)(0) = 0$$

$$\text{b) } (\sigma_i \wedge \sigma_j)(\mathcal{A}) \geq 0 \quad \forall \mathcal{A} \in \mathcal{J}$$

$$\text{c) } (\sigma_i \wedge \sigma_j)\left(\bigcup_k \mathcal{A}_k\right) = \sum_k (\sigma_i \wedge \sigma_j)(\mathcal{A}_k) \quad \forall \mathcal{A}_k, \mathcal{A}_l \in \mathcal{J}$$

$$\mathcal{A}_k \cap \mathcal{A}_l = 0 \quad \text{for } k \neq l.$$

Inference space: $(\mathcal{M} \times \mathcal{D}, \mathcal{P}; \wedge)$

Rozkład nie-informacyjny $\mu(\cdot)$

Pozostaje więc konstrukcja rozkładu nieinformacyjnego $\mu(\mathbf{x})$
Rozwiązaniem cząstkowym (?) jest zapostulowana przez
Gaussa (w innym kontekście) zasada “insufficient reasoning”

$$P(\mathcal{A}) \equiv \text{vol}(\mathcal{A}) \rightarrow \int_{\mathcal{A}} \mu(\mathbf{m}) \, d\mathbf{m} \equiv \int_{\mathcal{A}} dV$$

Jednakże jest ona słuszna tylko dla parametrów które są
nieograniczone ($\mathbf{x} \in R^N$)

Rozkład μ - paradoks Borela

Przykład $v \in R^+$ oraz $s = 1/v$

$$P(\mathcal{A}_v) \equiv \text{vol}(\mathcal{A}_v) \rightarrow \int_{\mathcal{A}} \mu_v(v) dv = \int_{\mathcal{A}} dv \rightarrow \mu_v(v) = \text{const.}$$

$$P(\mathcal{A}_s) \equiv \text{vol}(\mathcal{A}_s) \rightarrow \int_{\mathcal{A}_s} \mu_s(s) ds = \int_{\mathcal{A}_s} ds \rightarrow \mu_s(s) = \text{const.}$$

ale

$$v = \frac{1}{s} \implies \mu_s(s) = \mu_v(v) \left| \frac{dv}{ds} \right|_{v=\frac{1}{s}} = \frac{1}{s^2}$$

Cóż zrobić w takim przypadku

Zanurzyć przestrzeń \mathcal{J} w odpowiedniej R^N i znaleźć $\mu_{\mathcal{J}}$ indukowane z R^N

Przykład: $\mathbf{x} \in \mathcal{M} = [a, b]$

Zanuzenie \mathcal{M} można zrealizować np.

$$x \rightarrow y = g(x) = \log \left(\frac{x - a}{b - x} \right)$$

Indukowany z R^N rozkład nie-informatyczny

$$\mu(x) = \frac{b - a}{(x - a)(b - x)}$$

Niejednoznaczność

Czy $\mu(\mathbf{x})$ jest jednoznaczne?

$$x \rightarrow y = g(x); \quad x \in \mathcal{M}, \quad y \in R^1$$

$$\mu(x) = \left| \frac{dg}{dx} \right| \times [\mu(y) = \text{const}]$$

Np. dla $x \in R^+$

$$g(x) = \ln(x) \quad \rightarrow \quad \mu(x) = 1/x$$

$$g(x) = x - 1/x \quad \rightarrow \quad \mu(x) = \frac{(1+x^2)}{x^2}$$

Symetria (niezmienniczość) stanu non-informative

Rozważmy transformacje

$$x \rightarrow x' = T(x)$$

względem których $\mu(\cdot)$ jest niezmiennicze ($\mu \equiv \mu'$)

$$\mu[T(x)] \left| \frac{dT}{dx} \right| \equiv \mu(x)$$

Z drugiej strony wykonując zanurzenie $g(\cdot) \mathcal{J} \rightarrow R$ mamy

$$\mu(x) = \text{const} \left| \frac{dg}{dx} \right|$$

co daje warunek

$$\left| \frac{d}{dx} g[T(x)] \right| = \left| \frac{d}{dx} g(x) \right|$$

Symetria (niezmienniczość) stanu non-informative - przykład

dla dodatnich wielkości fizycznych $v \in R^+$

$$T(v) = \frac{1}{v}$$

warunek niezmienniczości

$$\mu(v^{-1}) = v^2 \mu(v)$$

prowadzi do równania na zanurzenie $g(\cdot)$

$$g(v^{-1}) = -g(v) + \text{const}$$

Symetria (niezmienniczość) stanu non-informative - przykład

z rozwiązaniem

$$g(v) = a_0 \ln(v) + \sum_{l=1}^{\infty} a_l \left(v^l - \frac{1}{v^l} \right)$$

oraz

$$\mu(v) = \frac{1}{v} + \sum_{l=0}^{\infty} a_l \left(v^l - \frac{1}{v^{l+2}} \right)$$

Spontaniczne łamanie symetrii !!!

Inference space

Przestrzeń “inference space”

$$\mathcal{S} = (\mathcal{J}, \mathcal{P}, \wedge)$$

Wiemy już, że (\mathcal{P}, \wedge) jest półgrupą abelową

W \mathcal{S} możemy wprowadzić metrykę Fishera a wtedy \mathcal{S} staje się przestrzenią pseudo-riemanowską.

Jakie własności ma przestrzeń \mathcal{S} ?

Thank you for your attention

