

# **Metody inwersji Bayesowskiej**

## **-L6-**

IGF PAN, 14.IV.2005

# Podejście optymalizacyjne i probabilistyczne podobieństwa i różnice (A)

Podejście optymalizacyjne (klasyczne)

$$\| \mathbf{G}(\mathbf{m}) - \mathbf{d}^{obs} \| + \lambda \| \mathbf{m} - \mathbf{m}^a \| = \min$$

- ❖ metoda nieliniowa
- ❖ fizyczny sens regularyzacji
- ❖ poszukujemy jednego (optymalnego) rozwiązania
- ❖ zaniedbane wszystkie błędy i niedokładności

# Podejście optymalizacyjne i probabilistyczne podobieństwa i różnice (B)

Podejście probabilistyczne (Bayesowskie)

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

- ❖ metoda nieliniowa
- ❖ fizyczny sens członu  $f(\mathbf{m})$
- ❖ przeszukiwana cała przestrzeń  $\mathbf{M}$  ( $\mathbf{M} \times \mathbf{D}$ )
- ❖ uwzględnienie wszystkich błędów i niedokładności

# Podójście optymalizacyjne i probabilistyczne podobieństwa i róźnice (C)

$$\|G(\mathbf{m}) - \mathbf{d}^{obs}\| + \lambda \|\mathbf{m} - \mathbf{m}^{apr}\| = \min$$

$$\mathbf{d}^{obs} = \mathbf{d}^{true} + \epsilon_{obs}$$

$$\mathbf{m}^{apr} = \mathbf{m}^{true} + \epsilon_{apr}$$

$$\mathbf{d}^{th} = G(\mathbf{m}) + \epsilon_{th}$$

$$\|G(\mathbf{m}) - \mathbf{d}^{true} + \epsilon_{obs} + \epsilon_{th}\| + \lambda \|\mathbf{m} - \mathbf{m}^{apr} + \epsilon_{apr}\| = \min$$

# Podejście optymalizacyjne i probabilistyczne podobieństwa i różnice (D)

Rozwiązania dla różnych realizacji  $\epsilon_{obs}$ ,  $\epsilon_{th}$   $\epsilon_{apr}$ :

$$\mathbf{m}^{(1)}, \mathbf{m}^{(2)}, \dots \mathbf{m}^{(K)}$$

Pozwala utworzyć histogram rozwiązań:

$$h(\mathbf{m}) = \frac{N^\alpha(\mathbf{m})}{N^K}$$

który jest przybliżeniem  $\sigma(\mathbf{m})$

Technika nieefektywna w przypadkach wielowymiarowych!!!

# Rozwiązywanie zagadnień odwrotnych

## co musimy wiedzieć?

| wielkości                                          | optymalizacyjne | probabilistyczne        |
|----------------------------------------------------|-----------------|-------------------------|
| dane pomiarowe                                     | tak             | tak(zależy od problemu) |
| związek $\mathbf{d}^{th} = \mathbf{G}(\mathbf{m})$ | tak             | tak(zależy od problemu) |
| $\mathbf{m}^{apr}$                                 | tak             | nie                     |
| wsp. tłumienia                                     | tak             | nie                     |
| normę (y)                                          | tak             | nie                     |
| statystyka błędów                                  | nie             | tak                     |

## Normy ...

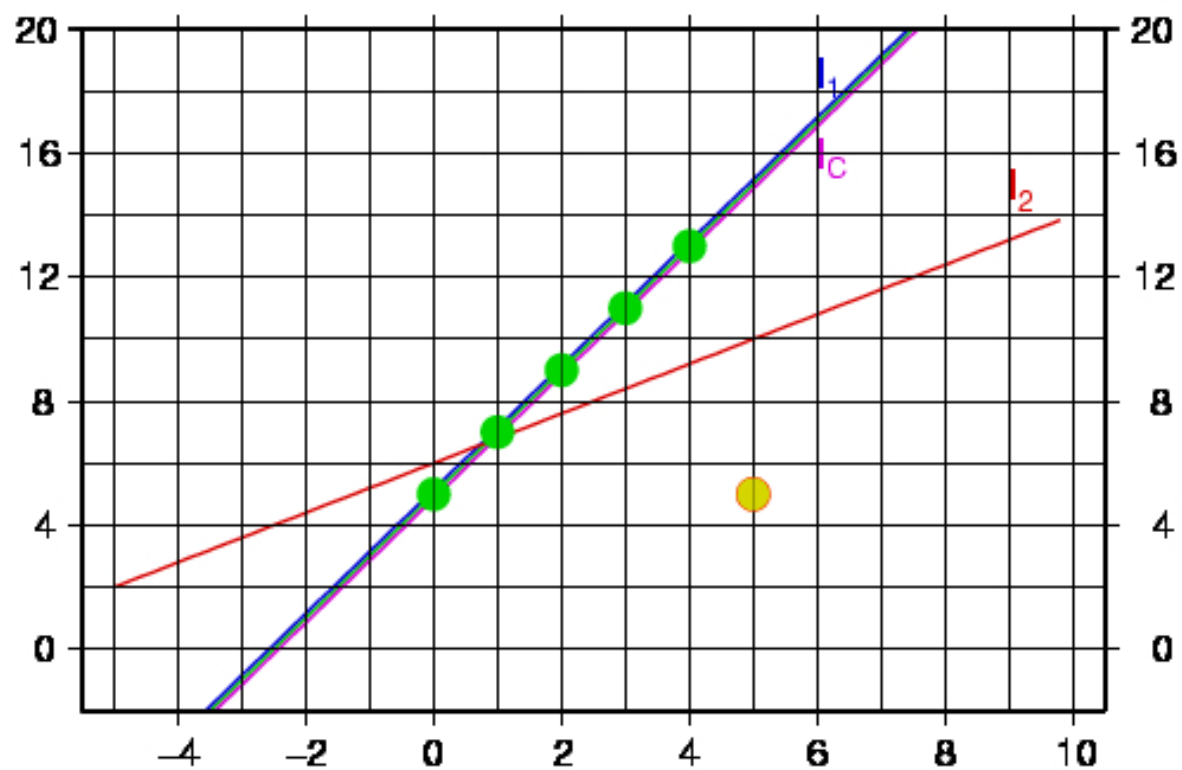
$$\|\mathbf{G}(\mathbf{m}) - \mathbf{d}^{obs}\|_D + \lambda \|\mathbf{m} - \mathbf{m}^{apr}\|_M = \min$$

---

$$\sigma(\mathbf{m}) = f(\mathbf{m}) \exp(-\|\mathbf{G}(\mathbf{m}) - \mathbf{d}^{obs}\|_D)$$

## “krzepkość” norm (ang. robustnes)

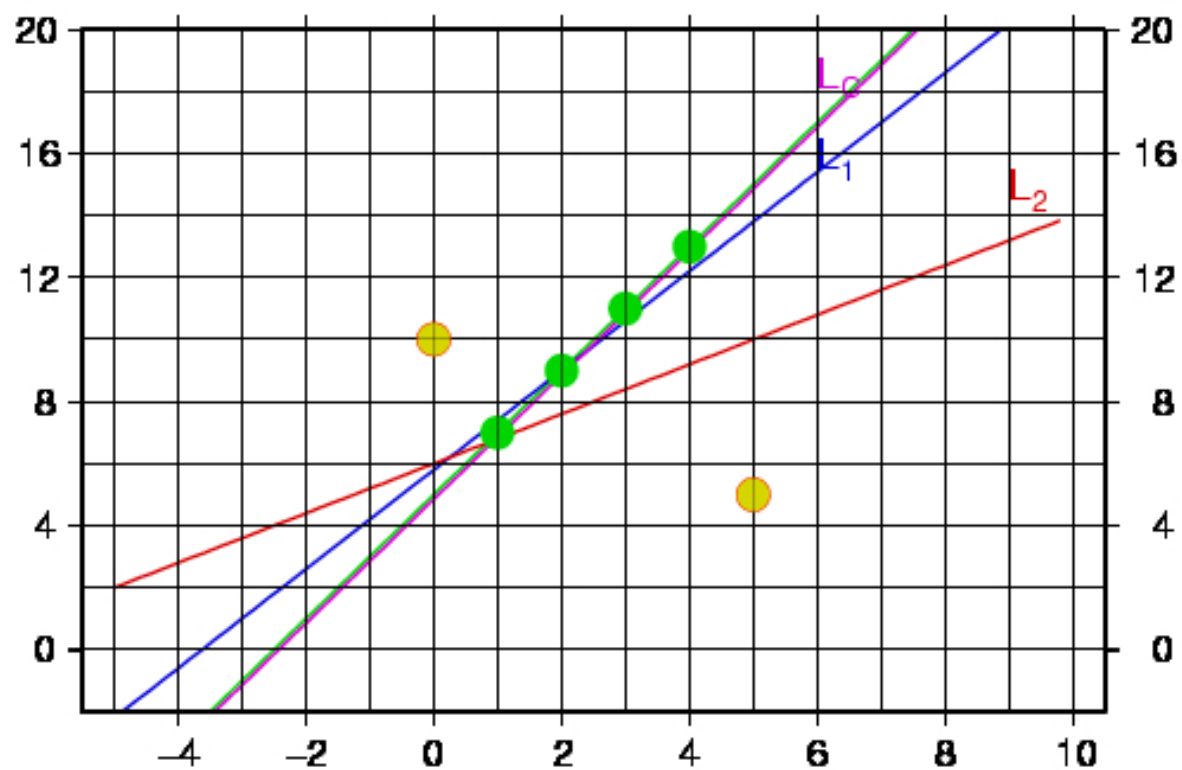
1 zaburzenie





## “krzepkość” norm (ang. robustnes)

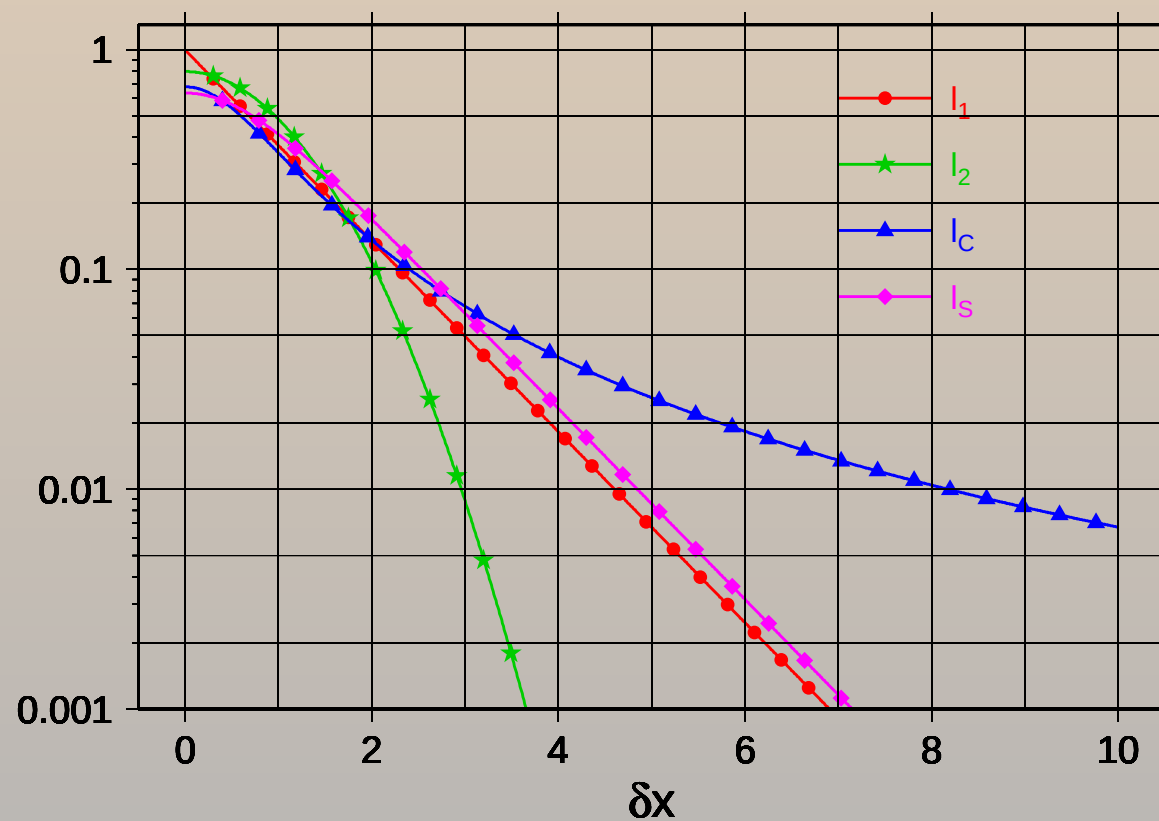
2 zaburzenia



# Generowane rozkłady pdf

$$\sigma(x) = a \exp(-||\delta x||)$$

Probability density functions



# Badanie rozkładu *a posteriori*

1. szukanie modelu najbardziej prawdopodobnego
  2. estymatory punktowe
  3. próbkowanie rozkładu *a posteriori*  
(rozkłady warunkowe i brzegowe)
-

# Model najbardziej prawdopodobny - optymalizacja

## ❖ szukanie modelu najbardziej prawdopodobnego

$$\mathbf{m}^{ml} : \quad \sigma(\mathbf{m}^{ml}) = \max$$

★ metody gradientowe

★ metody Monte Carlo

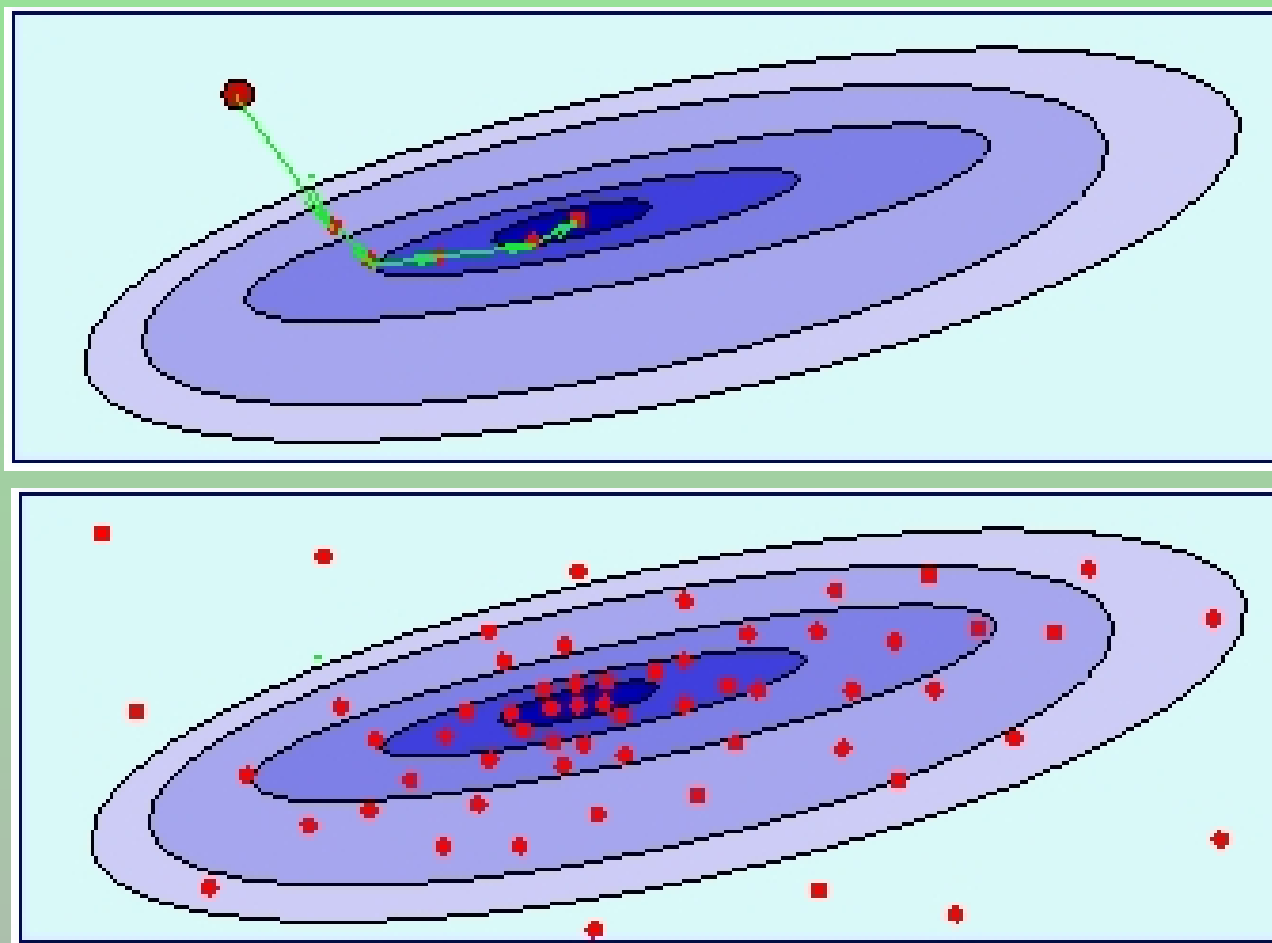
➡ Symulowane wyżarzanie

➡ Algorytm Genetyczny

---

technika równoważna podejściu optymalizacyjnemu

# Optymalizacja gradientowa i Monte Carlo



## Metoda: preconditioned steepest descent

1.  $\mathbf{m}_0$  - dowolny,  $\hat{\mathbf{S}}_0 \approx (\mathbf{I} + \mathbf{C}_M \mathbf{G}^t \mathbf{C}_D^{-1} \mathbf{G})^{-1}$
  2.  $\gamma_n = \mathbf{C}_M \mathbf{G}^t \mathbf{C}_D^{-1} (\mathbf{G} \mathbf{m}_n - \mathbf{d}^{obs}) + (\mathbf{m}_n - \mathbf{m}^{apr})$
  3.  $\phi_n = \hat{\mathbf{S}}_0 \gamma_n$
  4.  $\mathbf{b}_n = \mathbf{G} \phi_n$
  5.  $\mu_n = \frac{\gamma_n^t \mathbf{C}_M^{-1} \phi_n}{\phi_n^t \mathbf{C}_M^{-1} \phi_n + \mathbf{b}_n^t \mathbf{C}_D \mathbf{b}_n}$
  6.  $\mathbf{m}_{n+1} = \mathbf{m}_n - \mu_n \phi_n$
-

## Badanie rozkładu *a posteriori*

### ❖ punktowe estymatory (statystyczne)

- ★ model średni
- ★ model typu “mediana”
- ★ macierz kowariancji

❖ średni:  $\mathbf{m}^{avr} = \int_M \mathbf{m} \sigma(\mathbf{m}) d\mathbf{m}$

❖ mediana:  $\mathbf{m}^{med}$ :  $P(\mathbf{m} < \mathbf{m}^{med}) = P(\mathbf{m} > \mathbf{m}^{med})$

❖ macierz kowariancji:

$$C_{ij} = \int_M (m_i - m_i^{avr})(m_j - m_j^{avr}) \sigma(\mathbf{m}) d\mathbf{m}$$

# Badanie rozkładu *a posteriori*

## ❖ **Próbkowanie rozkładu *a posteriori***

### ★ próbkowanie “geometryczne”

➡ sieć regularna

➡ sieć adaptacyjna (near neighborhood algorithm)

### ★ próbkowanie stochastyczne

### ★ rozkłady brzegowe i warunkowe

---



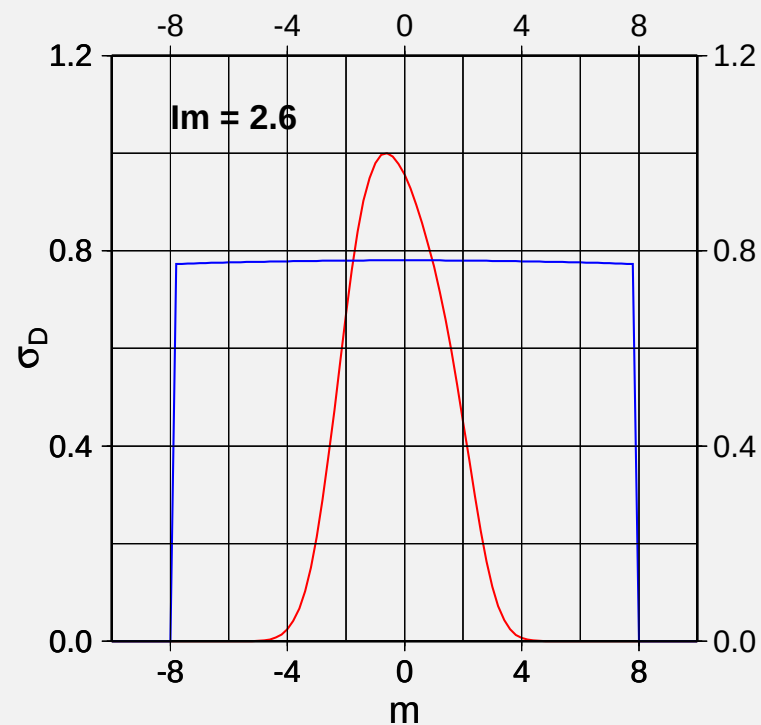
# Próbkowanie geometryczne

$$\sigma(\mathbf{m}) \Rightarrow \sigma_{ijk\dots} = \sigma(m_i^1, m_j^2, m_k^3, \dots)$$

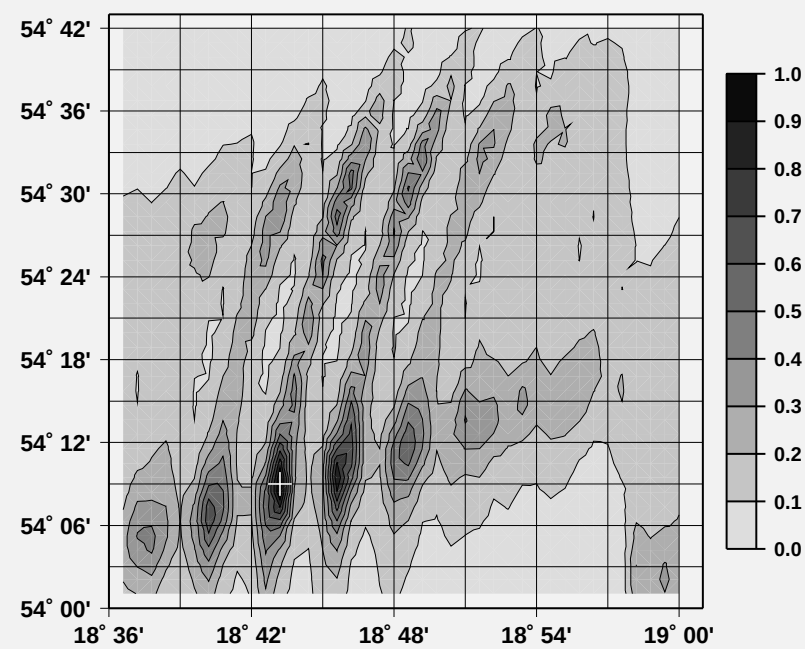
- ❖ **Metoda bardzo ogólna**
- ❖ **Tylko gdy mała liczba wymiarów  $N < 6$**
- ❖ **Bardzo czasochłonne obliczenia**

# Próbkowanie na sieci regularnej

$$N_m = 1$$



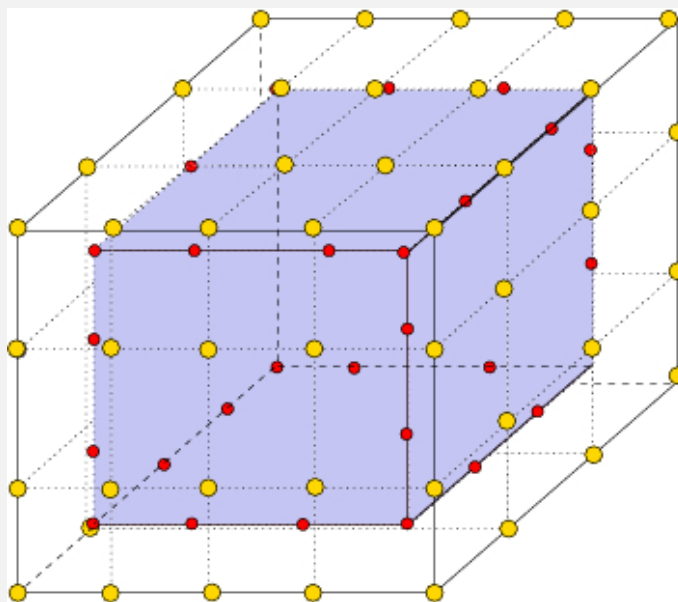
$$N_m = 2$$



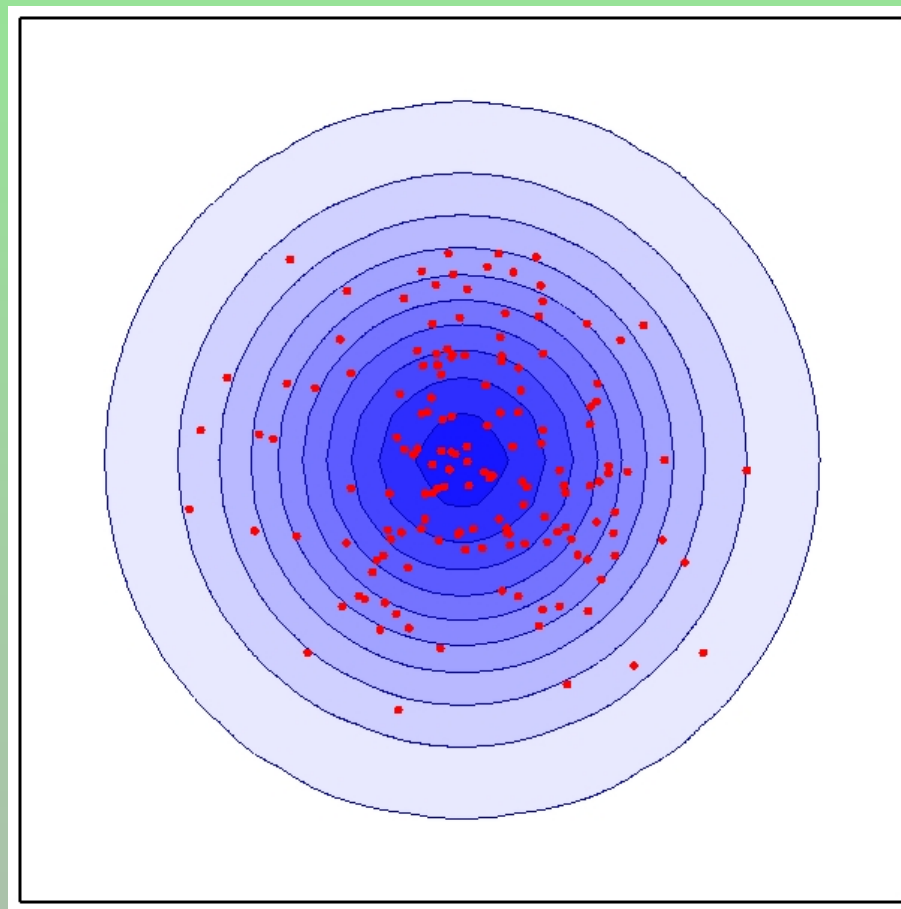
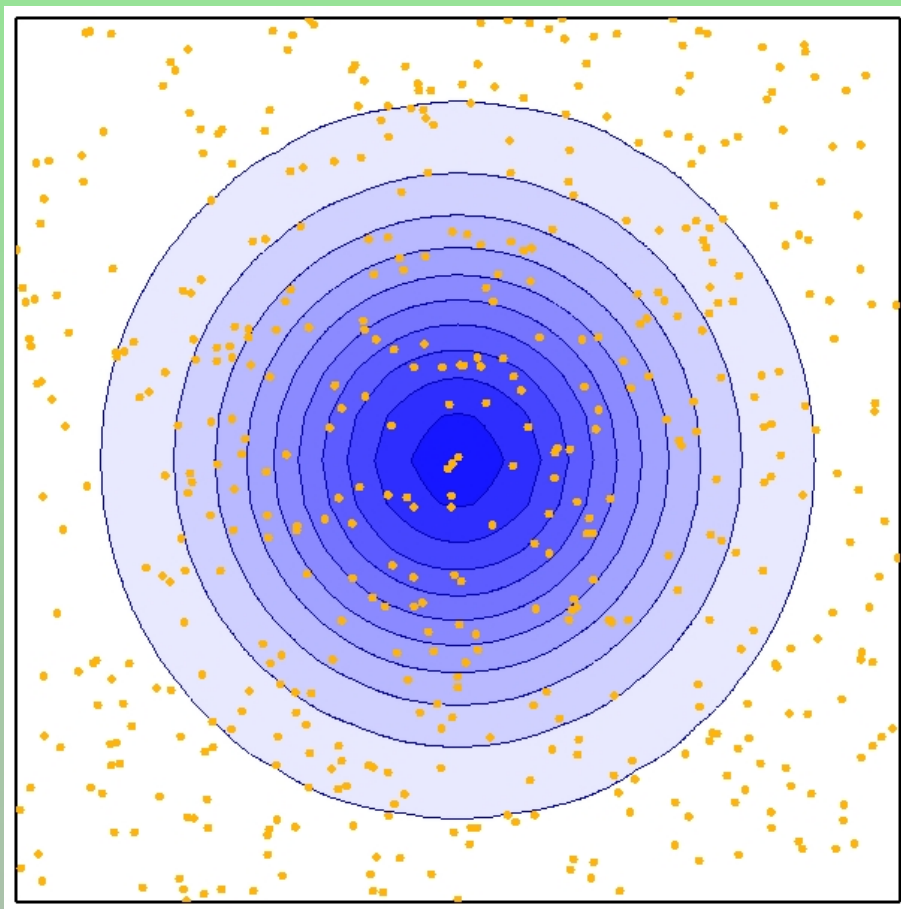
# Próbkowanie geometryczne w wielu wymiarach

**Próbkowanie niejednorodne:**

$$\frac{N_V}{N} = \left( \frac{p-2}{p} \right)^N \underset{p \gg 2}{\approx} e^{-2N/p} \xrightarrow{N \rightarrow \infty} 0$$



# Próbkowanie Monte Carlo



# Efektywność inwersji

$$\mathbf{m}^{apr} \Longrightarrow \mathbf{m}^{pos}$$

## Ocena uzyskanych nowych informacji

$$I^{ef} = \int_M \sigma^{pos}(\mathbf{m}) \log \left( \frac{\sigma^{pos}(\mathbf{m})}{\sigma^{apr}(\mathbf{m})} \right) d\mathbf{m}$$

## Analiza residuów

$$\mathbf{r}^{pos} = \mathbf{d}^{obs} - \mathbf{G}(\mathbf{m}^{pos})$$

$$\mathbf{r}^{apr} = \mathbf{d}^{obs} - \mathbf{G}(\mathbf{m}^{apr})$$

# Podsumowanie teorii

❖ zagadnienie odwrotne - uogólniony pomiar

❖ trzy różne podejścia

1. podejście algebraiczne
2. podejście optymalizacyjne
3. podejście probabilistyczne (Bayesowskie)

❖ rozwiązanie - **rozkład prawdopodobieństwa** a nie pojedynczy “optymalny model”

❖ konieczność efektywnego próbkowania wielowymiarowych rozkładów prawdopodobieństwa