Advanced statistical methods and Bayesian inference in scientific research

Lecture 7

W. Dębski

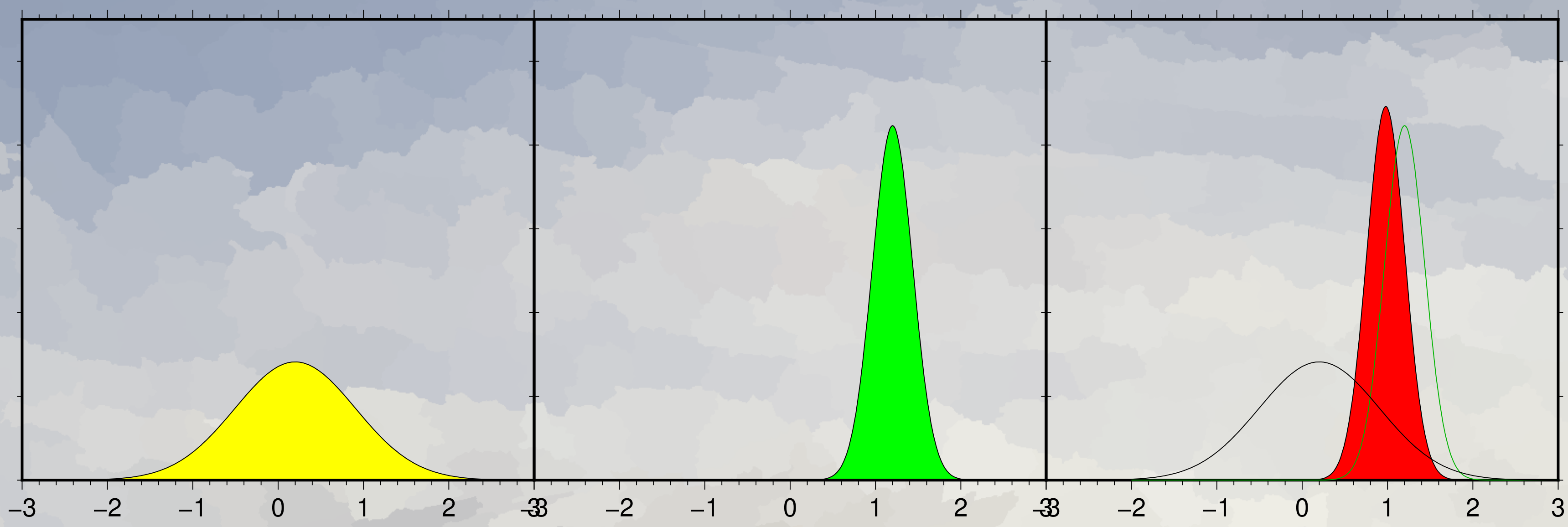
13.05.2024

Bayesian Inference

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

Essentially, $\mu(\cdot)$ can be arbitrary but usually is taken as *volumetric pdf*

Bayesian Inference



Bayesian inference (inversion)

- random variable X is described by $\rho(x)$
- we perform a measurement of Y
- measurement errors are characterized by $\psi(y - y_o)$ distribution
- we know that X and Y are related as $Y \; = \; G(X) + \epsilon$
  and the relation is subjected to errors ϵ described by

$$\zeta(X, Y) \; = \; \zeta(Y - G(X))$$

Question:

how performed measurement constraints (update) knowledge of X ?

Bayesian inference (inversion)

New “vectorized” random variable

$$X, Y \Longrightarrow Z := \begin{pmatrix} X \\ Y \end{pmatrix}$$

$$\mathcal{R} \times \mathcal{R} \Longrightarrow \mathcal{R}^2$$

$$\begin{aligned} \rho(x) &\rightarrow \rho'(z) = \rho(x)\mu(y) \\ \psi(y - y_o) &\rightarrow \psi'(z) = \psi(y - y_o)\mu(x) \\ \zeta(x, y) &\rightarrow \zeta'(z) \end{aligned}$$

Bayesian inference (inversion)

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

$$p(z) \ = \ a\,\frac{\psi'(z)\zeta'(z)}{\mu(z)}$$

$$\sigma(z) \ = \ a\,\frac{p(z)\rho'(z)}{\mu(z)}$$

$$\sigma(z) \ = \ \frac{1}{Z}\,\frac{\zeta(z)\,\psi'(z)\,\rho'(z)}{\mu^2(z)}$$

Bayesian inference (inversion)

Taking marginal integrals

$$\sigma(x) = \int_Y \sigma(z)\mathrm{d}y$$

$$\sigma(x) = \frac{1}{Z}\rho(x)\int_Y \psi(y-y_o)\zeta(y-G(x))\mathrm{d}y$$

$$Z = \int_X\int_Y \rho(x)\psi(y-y_o)\zeta(y-G(x))\mathrm{d}y\mathrm{d}x$$

Bayesian inference (inversion)

If the relation between X and Y is exactly known

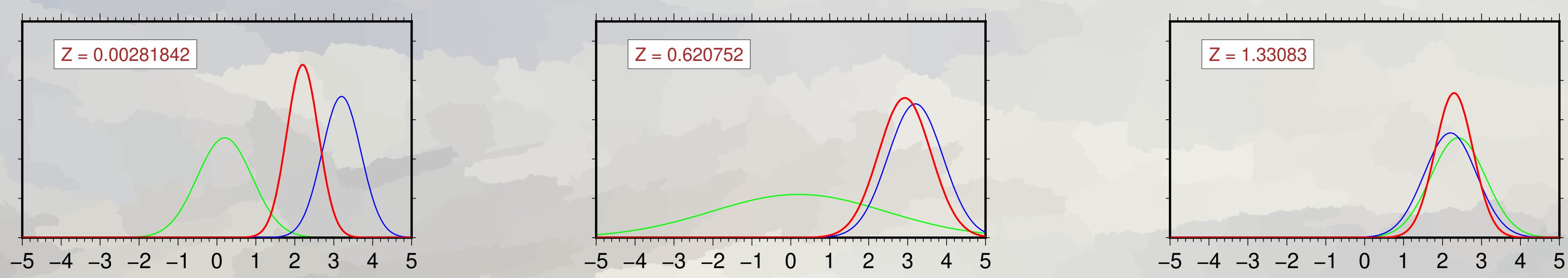
$$\zeta(y - G(x)) = \delta(y - G(x))$$

$$\sigma(x) = \frac{1}{Z}\,\rho(x)\,\psi(y_o - G(x))$$

$$Z = \int_X \rho(x)\psi(y_o - G(x))\mathrm{d}x$$

Evidence

- $\rho(x)$ - green
- $\mathcal{L}(x)$ - blue
- $\sigma(x) = \rho(x)\mathcal{L}(x)$ - red



Exploring pdf distribution

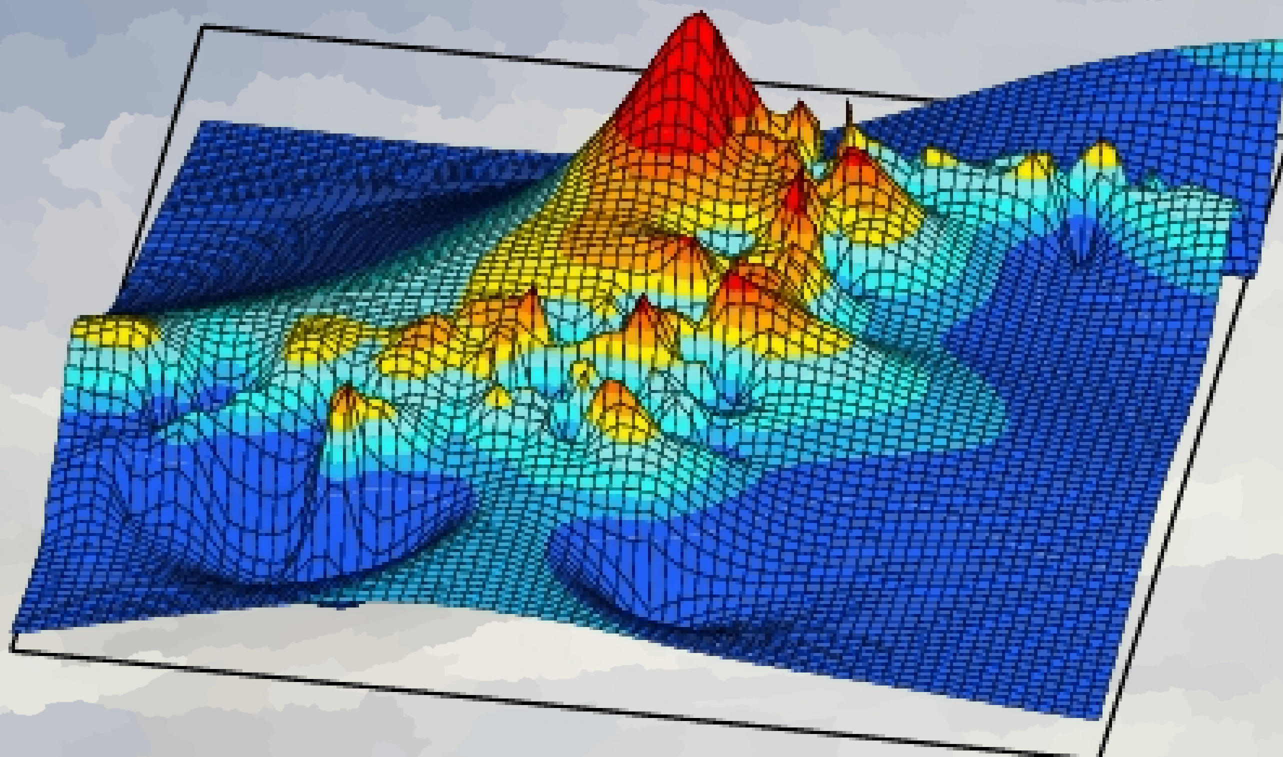
Stochastic inference about random variable X described by given probability distribution $\sigma(X)$ means, from technical point of view, that we need to explore properties of $\sigma(X)$

This can be carried out either analytically if $\sigma(X)$ is given by close analytical formula, like e.g.

$$\sigma(X) \; = \; \frac{1}{Z} \exp\left(-\mathbf{X}^T \cdot C^{-1} \cdot \mathbf{X}\right)$$

Exploring pdf distribution

or numerically in most of realistic cases.



Exploring pdf distribution

Depending on posed question (numerical) exploration of $\sigma(X)$ can generally be reformulated either as an optimization, sampling, or integration task

- searching for maximum of $\sigma(\mathbf{x})$
- calculate expected values of X (average,dispersion,...)
- calculate marginal distributions (sampling) $\sigma(\mathbf{x})$

Efficient methods of calculation multi-dimensional integrals needed !

Marginal *a posteriori* distribution/Sampling

- 1D marginals

$$\sigma_i(x_i) = \int_{\mathbf{x}\neq x_i} \sigma(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

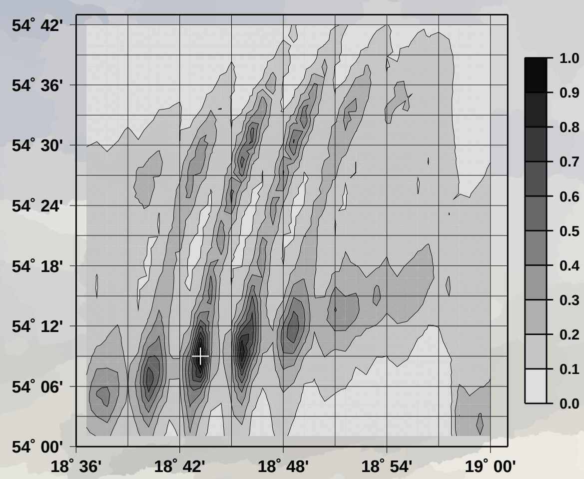
- 2D marginals

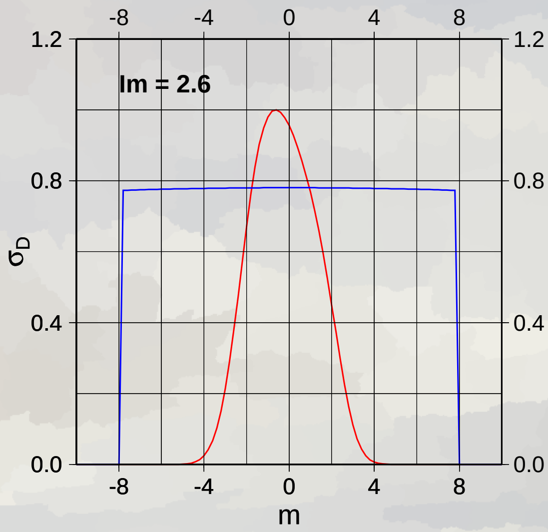
$$\sigma_{ij}(x_i, x_j) = \int_{\mathbf{x}\neq x_i, x_j} \sigma(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

- higher dimension marginals

Marginals

$$\sigma(\mathbf{r}) \;=\; k\exp\left(-\sum_{j=1}^{N_obs}(t_j^{obs} - t^{synth}(\mathbf{r}))^2\right)$$





Sampling *a posteriori* distribution

- geometric sampling - grid search
- stochastic (Monte Carlo sampling)

$$\sigma(\mathbf{x}) \Rightarrow \sigma_{i,j,k,\ldots} = \sigma(x_i, x_j, x_k, \ldots)$$

- very general: no limitation on $\sigma(X)$
- allows calculation maxima, moments, marginals (plots)
- only for small dimensional problem
- non-uniform sampling ...

Geometric Sampling - calculation time

N - dimensional random variable [100]

$$X = (X_1, X_2, \cdots, X_N)$$

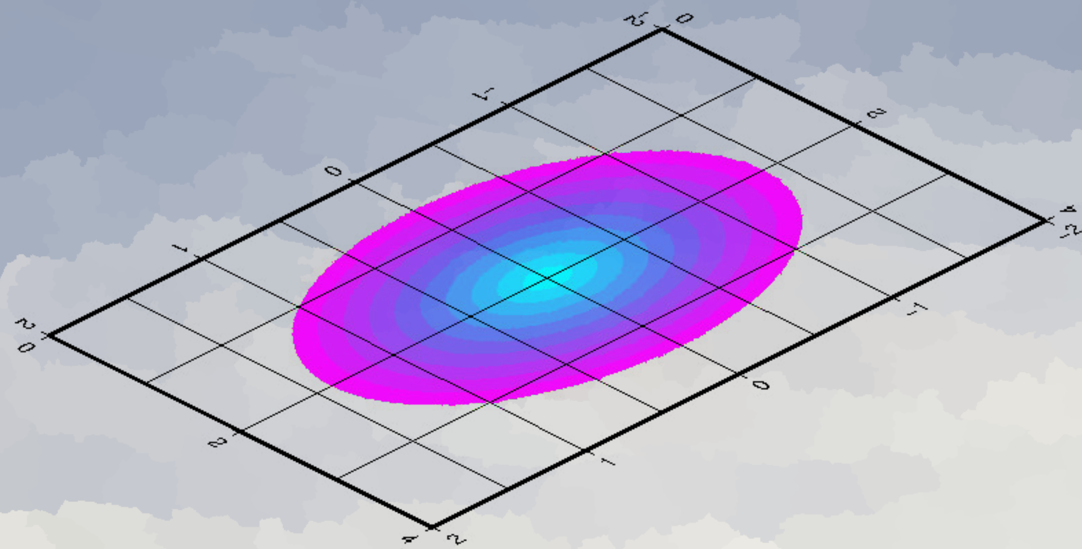
K - number of grid nodes for each dimension [10^3]

dT - calculation time for single grid node [10^{-9} sec]

Total calculation time :

$$T = K^N dt \; : \qquad \left(10^3\right)^{100} \cdot 10^{-9} = 10^{297}\text{sec}$$

Geometric Sampling - uniformity

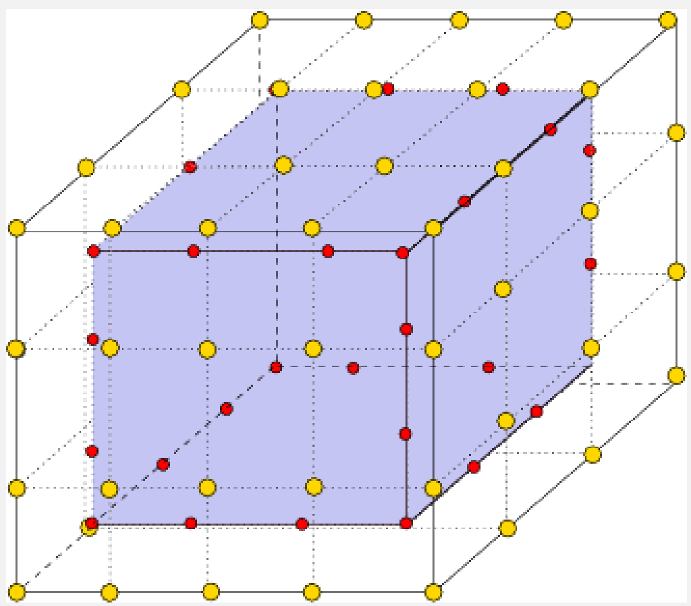


To sample $\sigma(x)$ we choose a range $[X_{min}, X_{max}]$ which is sufficiently large to encapsulate a region where $\sigma(x)$ is non-vanishing but small enough to avoid sampling in regions where $\sigma(x)$ is vanishing

Geometric sampling in multi-dimensional spaces

Non-uniform sampling:

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

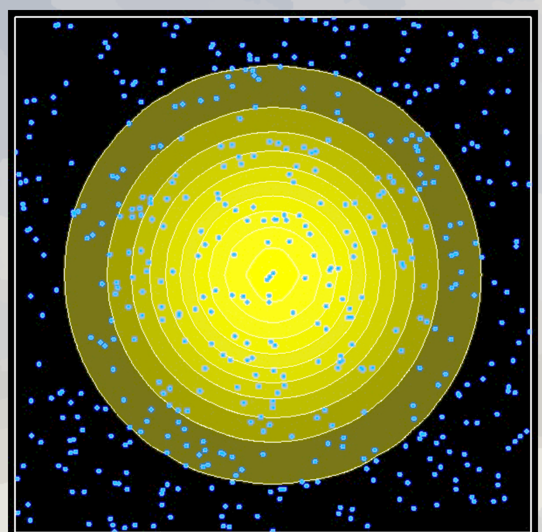


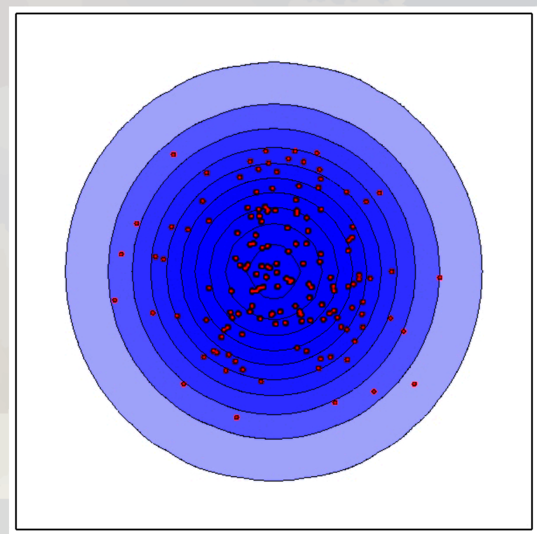
Regular grid sampling - curse of dimensionality problem

N/p	10	100	1000
2	0.64	0.96	0.996
3	0.51	0.94	0.994
5	0.33	0.90	0.990
10	0.10	0.82	0.980
100	$2.0\,10^{-10}$	0.13	0.818
1000	$1.2\,10^{-97}$	$1.6\,10^{-11}$	0.135
10000	—	$1.8\,10^{-88}$	$2\,10^{-9}$

Sampling *a posteriori* distribution

- geometric sampling - grid search
- stochastic (Monte Carlo sampling)





See you May, 20th