Lernen à la Bayes

—Von der Bildrekonstruktion zur inversen Quantentheorie —

Teil 3: Bildrekonstruktion, Mischungsmodelle und Hyperparameter

Jörg Lemm

Institut für Theoretische Physik
Universität Münster

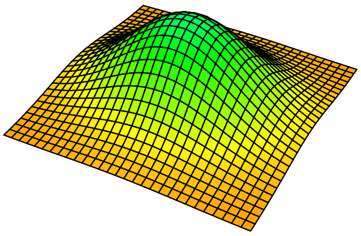
lemm@uni-muenster.de

http://pauli.uni-muenster.de/~lemm

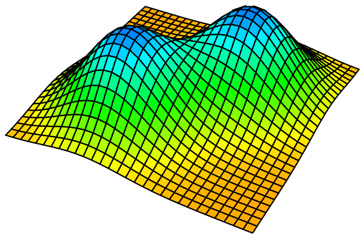
J.L., ICANN99; J.L., physics/9912005; J.L., Uhlig, Weiguny, quant-ph/0005122

Explizite Modellierung komplexer Prior mit Mischungsmodellen, Hyperparametern und Hyperfeldern

Gauß–Prozess

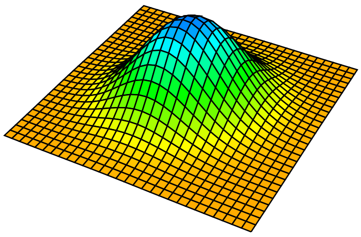


Mischung zweier Gauß–Prozesse

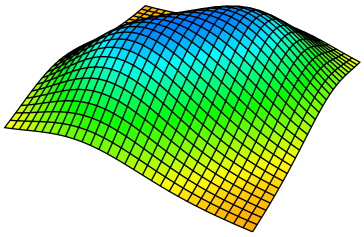


Tieftemperaturfall

Mischung zweier Gauß–Prozesse

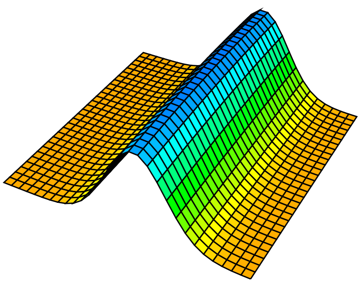


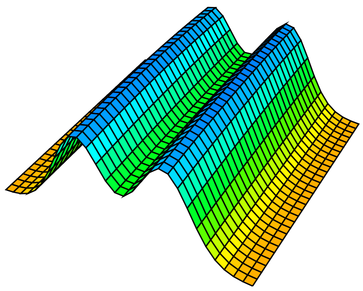
Hochtemperaturfall



beim „kritischen Punkt“

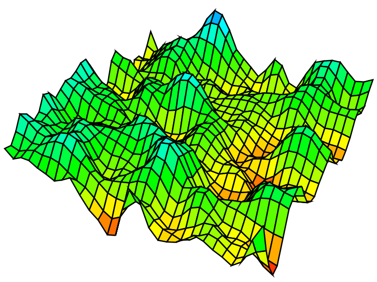
Hyperparameter θ oder Hyperfelder $\theta(x)$



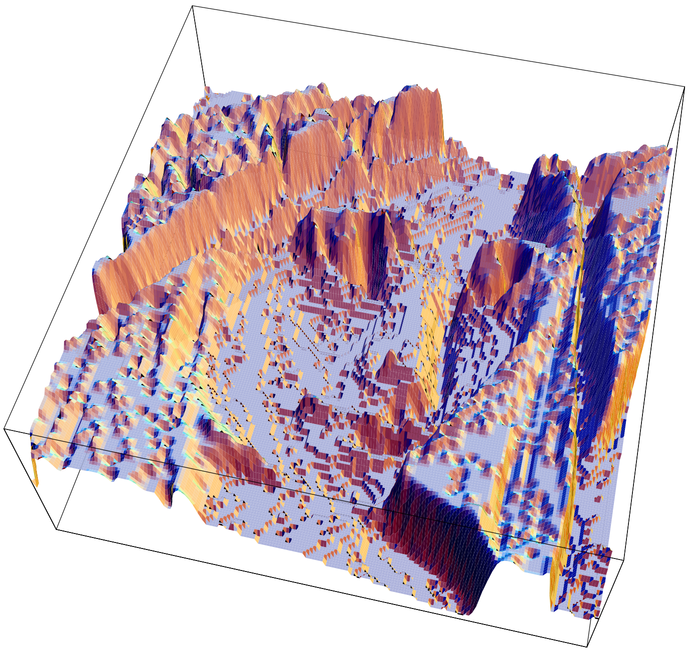


Mischung mit Hyperparametern/–feldern

Vgl. mit typischer Fehleroberfläche eines parametrisches Verfahrens
(z.B. eines neuronalen Netzes)

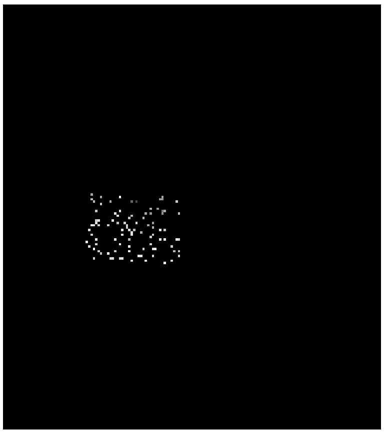


Beispiel einer zweidimensionalen Regressionsfunktion

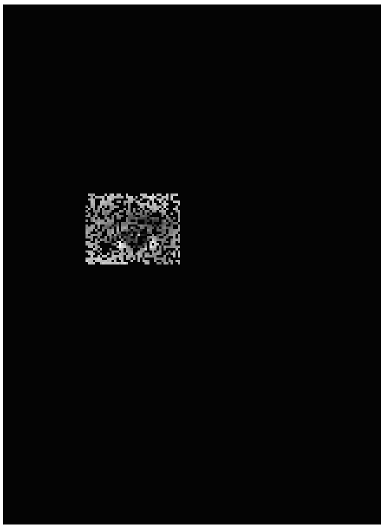


Bildrekonstruktion à la Bayes

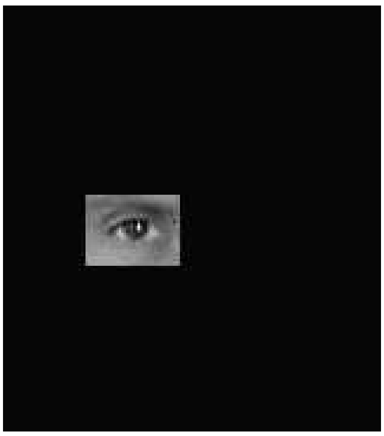
Beispieldaten:



100 Datenpunkte



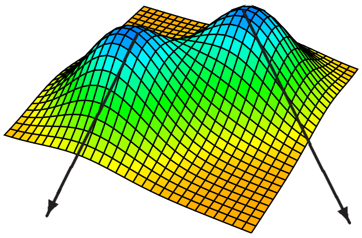
1000 Datenpunkte



10000 Datenpunkte

J.L., ICANN99 (physics/9911077) [J.L., EUFIT'98 + FNS'98]

Prior als Mischung von Gaußprozessen



t_1  „oder"  t_2

$$p(h) \propto \sum_{j=1}^{m} a_j \; e^{-\frac{1}{2}\left\langle h-t_j \middle| \mathbf{K}_j \middle| h-t_j \right\rangle}$$

Mittelwert der „predictive density"

$$\bar{y}(x) = \int dy\, y\, p(y|x,D) = \sum_j b_j\, \bar{t}_j$$

mit Komponentenlösung $\bar{t}_j = t_j + \mathbf{K}_j^{-1}\mathbf{K}_{t_D}(j)(t_D - t_j)$ und Mischungskoeffizient

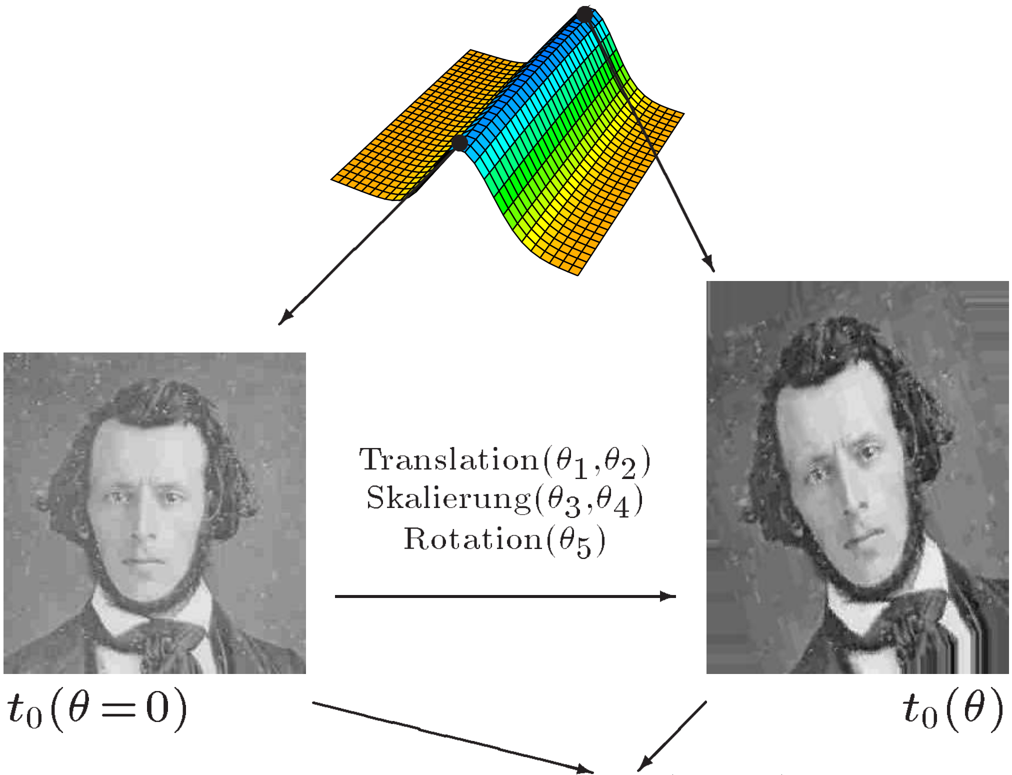
$$b_j = \frac{p(j)p(y_D|x_D,D_0,j)}{\sum_j p(j)p(y_D|x_D,j)},$$

$$p(y_D|x_D,j) = e^{-\frac{\beta}{2}\left\langle t_D-t_j \middle| \mathbf{K}_{t_D}(j) \middle| t_D-t_j \right\rangle + \frac{1}{2}\ln\det\left(\frac{\beta}{2\pi}\mathbf{K}_{t_D}(j)\right)}.$$

Die $\tilde{n} \times \tilde{n}$–Matrix $\mathbf{K}_{t_D}(j) = (\mathbf{K}_D^{-1} + \mathbf{K}_{j,DD}^{-1})^{-1}$ ist definiert im Raum der Daten und $\mathbf{K}_{j,DD}^{-1}$ ist die Projektion der Prior–Kovarianz $\mathbf{K}_j^{-1}$ in diesen Raum.

J.L., ICANN99 (physics/9911077) [J.L., EUFIT'98 + FNS'98]

Flexible Prior durch Hyperparameter



$$p(h) \propto \int d\theta \; a(\theta) \; e^{-\frac{1}{2}\left\langle h-t_0(\theta)\middle|\mathbf{K}_0(\theta)\middle|h-t_0(\theta)\right\rangle}$$

Mittelwert der „predictive density"(für Kombination mit Mischungmodell)

$$\bar{y}(x) = \int dy\, y\, p(y|x,D) = \sum_j \int d\theta\; b_j(\theta)\, \bar{t}_j(\theta)$$

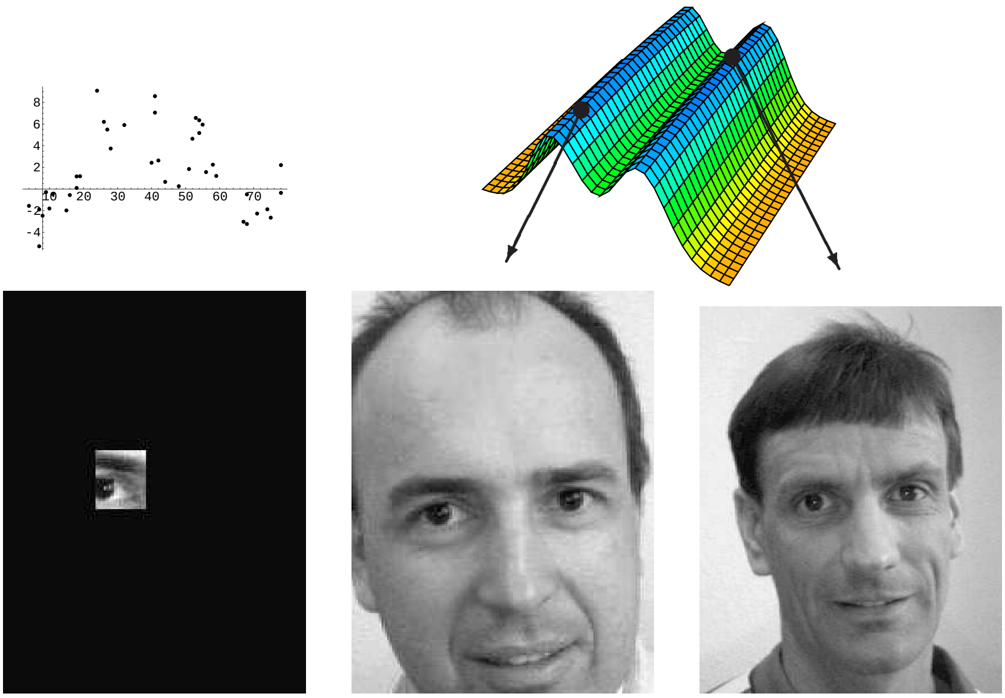
mit Mischungskoeffizienten

$$b_j(\theta) = \frac{p(\theta,j)\, p(y_D|x_D,\theta,j)}{\sum_j \int d\theta\, p(\theta,j)\, p(y_D|x_D,\theta,j)}$$

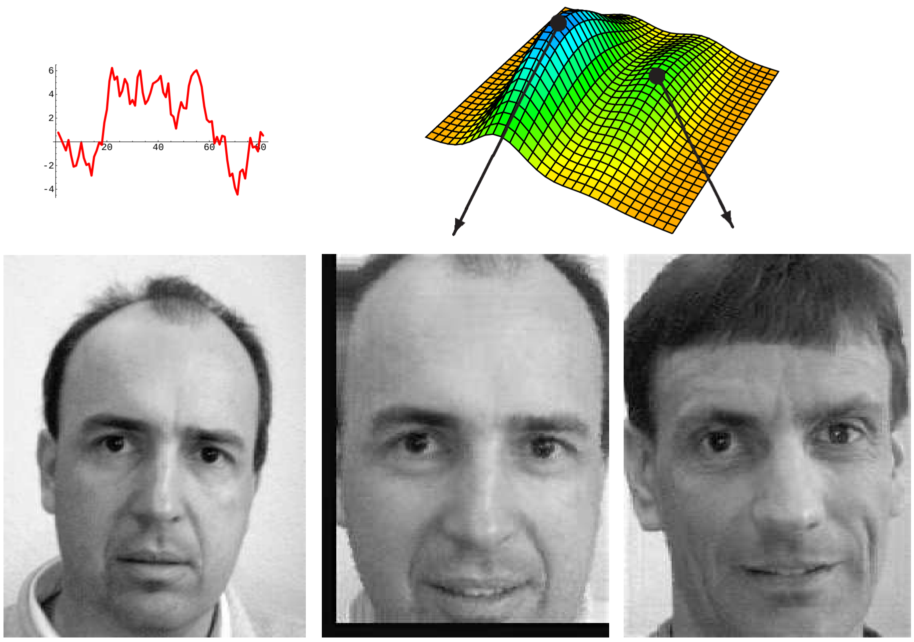
Die MAP–Stationaritätsgleichung für θ wird zu

$$\begin{aligned}
0 \quad = \quad & \sum_j \Bigg[\left\langle \frac{\partial t_j(\theta)}{\partial\theta} \middle| \mathbf{K}_{t_D}(\theta,j) \middle| t_j(\theta) - t_D \right\rangle \\
& + \frac{1}{2} \left\langle t_D - t_j(\theta) \middle| \frac{\partial \mathbf{K}_{t_D}(\theta,j)}{\partial\theta} \middle| t_D - t_j(\theta) \right\rangle \\
& - \mathrm{Tr}\left(\mathbf{K}_{t_D}^{-1}(\theta,j) \frac{\partial \mathbf{K}_{t_D}(\theta,j)}{\partial\theta} \right) - \frac{1}{p(\theta,j)} \frac{\partial p(\theta,j)}{\partial\theta} \Bigg]
\end{aligned}$$

Bildrekonstruktion: Tieftemperatur–Resultate

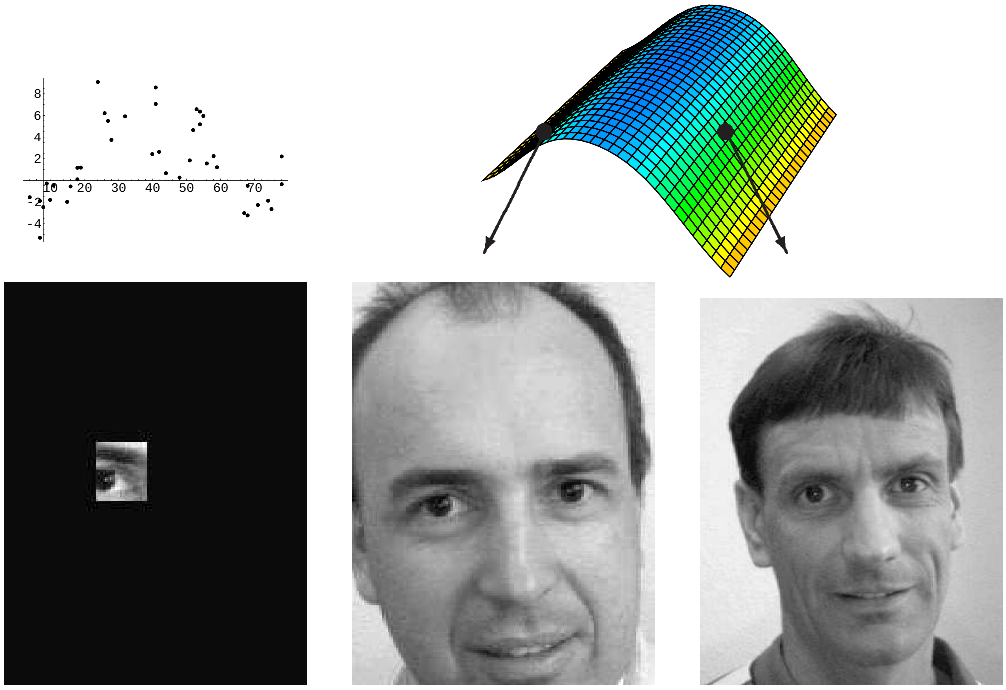


Daten t_1 t_2

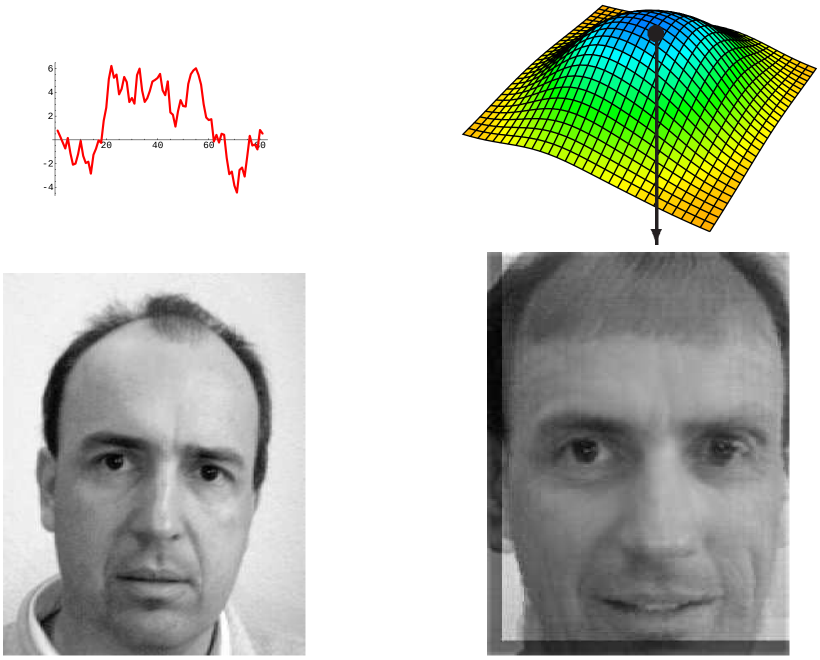


Original Rekonstruktion Lokales Maximum

Bildrekonstruktion: Hochtemperatur–Resultate



Daten t_1 t_2



Original

Hochtemperatur–
Rekonstruktion