Kriging and Radial Basis Functions

Hans Bruun Nielsen & Kristine Frisenfeldt Thuesen

DTU, IMM, Scientific Computing Section

- Introduction, definitions
- Kriging
- General RBFs
- Unified approach
- Examples
- Plans for future work

Introduction

$Y : \mathbb{R}^d \mapsto \mathbb{R}$ given by design points $\{(\underline{x}_i,\ y_i = Y(\underline{x}_i))\}_{i=1}^m$

Surrogate model: $s(\underline{x}) = \underline{\alpha}^T \underline{\phi}(\underline{\theta}, \underline{x}) + \underline{\beta}^T \underline{f}(\underline{x})$

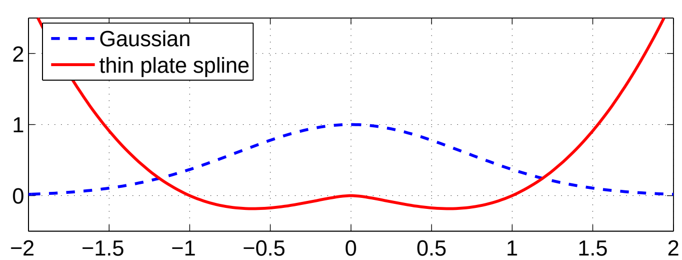
Regression (polynomial) part: $\underline{f}(\underline{x}) = (f_1(\underline{x}), \ldots, f_q(\underline{x}))^T, \quad f_j : \mathbb{R}^d \mapsto \mathbb{R}$

Correlation (radial) part: $\underline{\phi}(\underline{\theta}, \underline{x}) = (\phi_1(\underline{\theta}, \underline{x}), \ldots, \phi_m(\underline{\theta}, \underline{x}))^T$

where $\phi_i(\underline{\theta}, \underline{x}) = \phi(\|\Theta(\underline{x} - \underline{x}_i)\|_2), \quad \phi : \mathbb{R}_+ \mapsto \mathbb{R}$. scaling $\Theta = \mathrm{diag}(\theta_1, \ldots, \theta_d)$

Examples

Name	$\phi(r), \quad r \geq 0$
Gaussian	e^{-r^2}
inverse multiquadric	$(r^2+1)^{-1/2}$
multiquadric	$(r^2+1)^{1/2}$
thin plate spline	$r^2 \log r$



Interpolation

$$\Phi\underline{\alpha} + F\underline{\beta} \;=\; \underline{y}\,, \quad \Phi_{ij} \;=\; \phi_j(\underline{\theta}, \underline{x}_i), \quad F_{i,:} \;=\; \underline{f}(\underline{x}_i)^T$$

m equations with $m{+}q$ unknowns.

Both in Kriging and general RBF approach: supply with $F^T\underline{\alpha} = \underline{0}$

$$\begin{pmatrix} \Phi & F \\ F^\top & 0 \end{pmatrix} \begin{pmatrix} \underline{\alpha} \\ \underline{\beta} \end{pmatrix} \;=\; \begin{pmatrix} \underline{y} \\ \underline{0} \end{pmatrix} \qquad (*)$$

$\Phi \in \mathbb{R}^{m\times m}$ is symmetric. Full rank if the $\underline{x}_i$ are distinct.

$F \in \mathbb{R}^{m\times q}$, $q < m$, is assumed to have rank q.

$(*)$ has a unique solution.

• Kriging

$$Y(\underline{x}) \;=\; \underline{\beta}^T \underline{f}(\underline{x}) + \zeta(\underline{x})\,, \qquad \underline{y} \;=\; F\underline{\beta} + \underline{z}\,,\; z_i = \zeta(\underline{x}_i)$$

$\zeta(\underline{x})$ stochastic, $\mathrm{E}\big[\zeta(\underline{x})\big] \;=\; 0$, $\mathrm{E}\big[\zeta(\underline{x})^2\big] \;=\; \sigma^2$, $\mathrm{E}\big[z_i\,\zeta(\underline{x})\big] \;=\; \sigma^2\phi_i(\underline{\theta}, \underline{x})$, $\mathrm{E}\big[\underline{z}\;\underline{z}^T\big] \;=\; \sigma^2\,\Phi$

Assumptions $\phi(0) = 1$ and Φ is positive definite.

The approximation $s(\underline{x})$ is a linear combination of the y_i, $s(\underline{x}) \;=\; \underline{\gamma}(\underline{x})^T\underline{y}$

MSE: $\Omega^2(\underline{x}) \;=\; \mathrm{E}\big[(s(\underline{x}) - Y(\underline{x}))^2\big] \;=\; \sigma^2[1 - 2\underline{\gamma}(\underline{x})^T\underline{\phi}(\underline{\theta}, \underline{x}) + \underline{\gamma}(\underline{x})^T\Phi\,\underline{\gamma}(\underline{x})]$

Minimize Ω^2 under the constraint $F^T\underline{\gamma}(\underline{x}) - \underline{f}(\underline{x}) = \underline{0}$

Estimated variance $\widetilde{\sigma}^2 \;=\; \dfrac{1}{m}\,\big(\underline{y} - F\underline{\beta}\big)^T\,\Phi^{-1}\,\big(\underline{y} - F\underline{\beta}\big)$

Assume Gaussian process: $\underline{\theta}^* = \mathrm{argmin}_{\underline{\theta}}\big\{\det\Phi(\underline{\theta})\cdot\widetilde{\sigma}^{2m}(\underline{\theta})\big\}$

General RBFs

Φ is not necessarily definite, but Powell[1] shows that a number of popular RBFs satisfy

$$\operatorname{sign}\left(\underline{v}^T \Phi\, \underline{v}\right) \;=\; (-1)^{\mu}$$

for all $\underline{v} \in \mathbb{V}_{\mu},\ \underline{v} \neq \underline{0}$

$$\mathbb{V}_{\mu} \;=\; \{\underline{v} \in \mathbb{R}^m : \textstyle\sum_{i=1}^{m} v_i p(\underline{x}_i) = 0 \text{ for any } p \in \Pi_{\mu-1}\}$$

We assume that $\{f_j\}$ comprises a basis of $\Pi_{\mu-1}$. Then $\mathbb{V}_{\mu} \subseteq \mathcal{N}(F^T)$, the nullspace of F^T. An orthonormal basis of $\mathcal{N}$ can be found as N in the complete QR factorization of F,

$$F \;=\; \begin{pmatrix} Q & N \end{pmatrix} \begin{pmatrix} R \\ 0 \end{pmatrix} \;=\; QR$$

Name	$\phi(r),\quad r \geq 0$	μ
Gaussian	e^{-r^2}	0
inverse multiquadric	$(r^2+1)^{-1/2}$	0
linear	r	1
multiquadric	$(r^2+1)^{1/2}$	1
thin plate spline	$r^2 \log r$	2
cubic	r^3	2

[1] M.J.D. Powell: *5 lectures on radial basis functions.* Report IMM-REP-2005-03, IMM, DTU, 2005.

$$s(\underline{x}) \;=\; \underline{\alpha}^T \underline{\phi}(\underline{\theta}, \underline{x}) + \underline{\beta}^T \underline{f}(\underline{x})\,. \qquad \Phi\, \underline{\alpha} + F\, \underline{\beta} \;=\; \underline{y}$$

Seek $\underline{\alpha} \in \mathbb{V}_{\mu} \subseteq \mathcal{N}(F^T)$: $\quad \underline{\alpha} = N\, \underline{\alpha}_N$

$$N^T \Phi\, N\, \underline{\alpha}_N + N^T F\, \underline{\beta} \;=\; N^T \underline{y}\,, \qquad (-1)^{\mu}\, \widetilde{\Phi}\, \underline{\alpha}_N \;=\; \underline{y}_N$$

The matrix $\widetilde{\Phi} = (-1)^{\mu}\, N^T \Phi\, N$ is symmetric and positive definite.

The regression coefficients are found from

$$R\, \underline{\beta} \;=\; Q^T \left(\underline{y} - \Phi N \underline{\alpha}_N\right)$$

• Unified approach

$$\underline{a} \in \mathbb{R}^m. \quad \underline{a} \;=\; Q\underline{a}_R + N\underline{a}_N, \quad \underline{a}_R = Q^T\underline{a}, \quad \underline{a}_N = N^T\underline{a}$$

Assume that $\underline{y} \;=\; F\underline{\beta} + \underline{z}, \quad \underline{z} = N\underline{z}_N$

$$\mathrm{E}\big[\underline{z}\;\underline{z}^T\big] \;=\; \sigma^2 NAN^T, \quad A \;=\; D\,\widetilde{\Phi}\,D, \quad D \;=\; \mathrm{diag}(\widetilde{\Phi}_{ii}^{-1/2})$$

A is symmetric, positive definite, and $A_{ii} = 1$. Valid correlation matrix. $A = D\,C^TC\,D$

BLUE: $\min\limits_{\underline{v},gb} \|\underline{v}\|_2^2$ s.t. $F\underline{\beta} + NDC^T\underline{v} = \underline{y}$ Solution: $C^T\underline{v} = (-1)^\mu D^{-1}N^T\underline{y}$

$\underline{v} \in \mathbb{R}^{m-q}$ has variance $\sigma^2 I$. Estimate: $\overline{\sigma}^2 \;=\; \|\underline{v}\|_2^2/(m-q)$

Assume Gaussian process: $\underline{\theta}^* = \mathrm{argmin}_{\underline{\theta}}\Big\{\Gamma(\underline{\theta}) \equiv \det A(\underline{\theta}) \cdot \overline{\sigma}^{2(m-q)}(\underline{\theta})\Big\}$

Theorem: For $\phi(r) \in \{r,\; r^3,\; r^2\log r\},\; \omega \in \mathbb{R}_+$:

$$s(\omega\underline{\theta}, \underline{x}) \;=\; s(\underline{\theta}, \underline{x}), \qquad \Gamma(\omega\underline{\theta}) \;=\; \Gamma(\underline{\theta})$$

Remarks:

"Scalar" $\underline{\theta}$ has no effect.

"Full" $\underline{\theta}$: Finding $\underline{\theta}^*$ reduces to a problem in $\mathbb{R}_+^{d-1}$ instead of $\mathbb{R}_+^d$

Examples

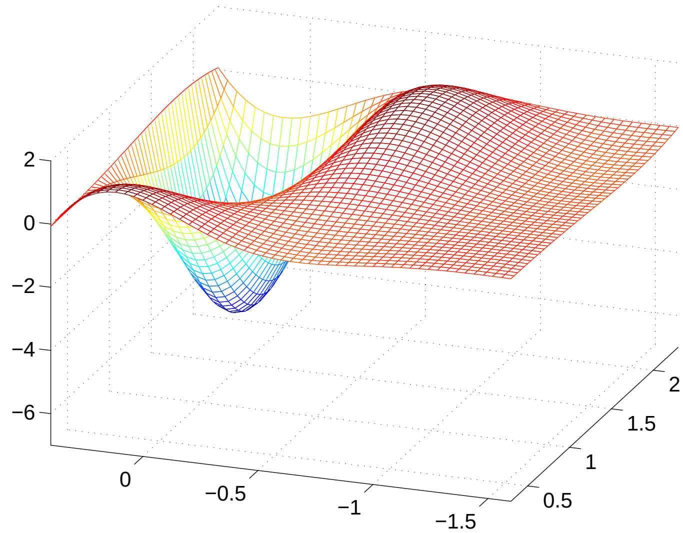
$$Y(\underline{x}) \;=\; \prod_{k=1}^{d} e^{k x_k}\,\cos(2k\,x_k)$$

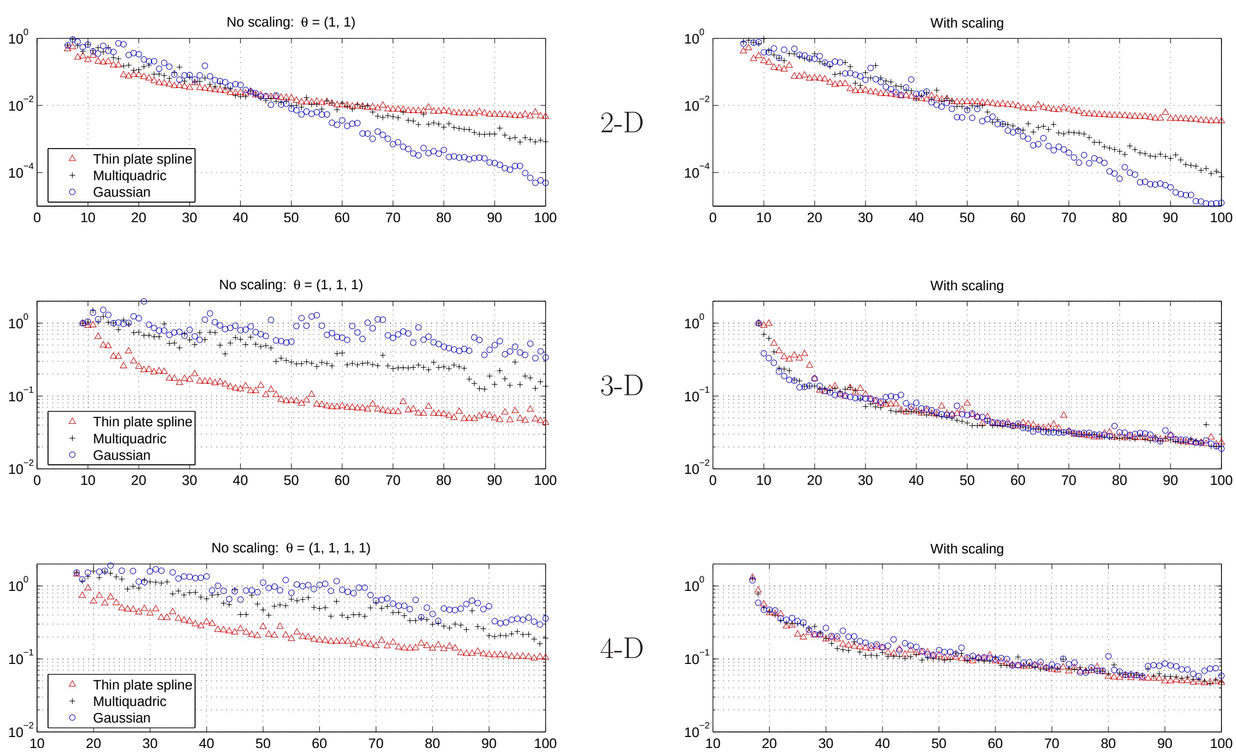
Educated guess: $\underline{\theta}^* = (1, 2, \ldots, d)^T$

Test regions $\mathcal{T}$ of sizes $2^2,\; 2^3,\; 0.5^4$

Successively insert new design point at

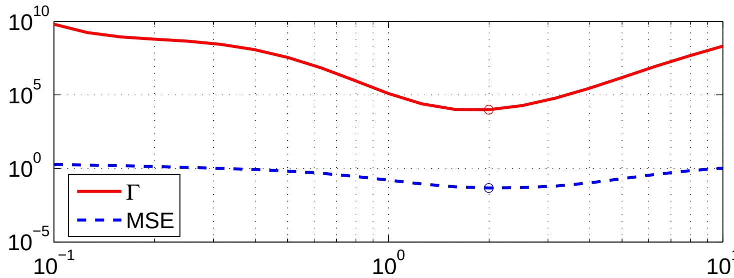
$$\mathrm{argmax}_{\underline{x}\in\mathcal{T}} |s(\underline{x}) - Y(\underline{x})|$$





Data: First 25 points from Gaussian example.

Use thin plate spline. $\underline{\theta}^* = (1,\ \theta_2^*)^T$

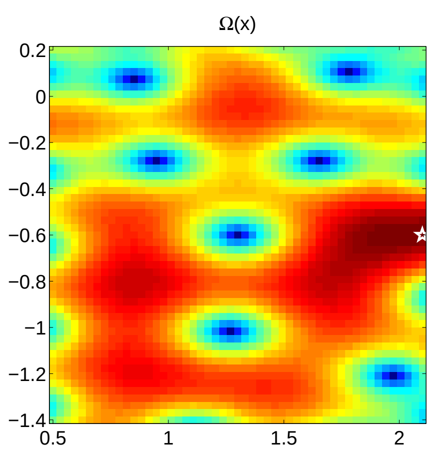


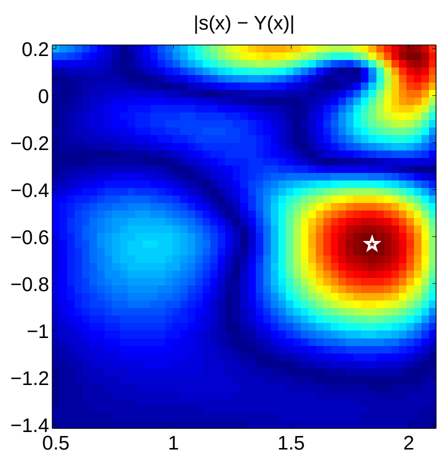
$\underline{\theta}^* = (1,\ 2)^T$ (as expected)

MSE: under progress.

Assume $\phi(0) \in \{0, 1\}$.

$$\Omega^2(\underline{x}) = (-1)^\mu \sigma^2 \left\{\phi(0) + \underline{\gamma}(\underline{x})^T \left[\Phi\underline{\gamma}(\underline{x}) - 2\underline{\phi}(\underline{\theta}, \underline{x})\right]\right\} \qquad (?)$$





Plans for future work

- Verifying (improving ?) the expression for $\Omega^2(\underline{x})$
- Large scale: can explicit computation of N be avoided ?
- Update the MATLAB DACE package[2] to allow for general RBFs

[2] See http://www2.imm.dtu.dk/~hbn/dace/