University of Antwerp
Faculty of Applied Engineering

Constraints on GP Predictions

Not so normal after all

dr. Ivan De Boi

8 September 2025

InViLab

Camera system development

Automation

Prototype development

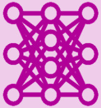
Camera integration

Advanced camera technologies

- UV → LWIR
- Multi- and hyperspectral
- 3D imaging
- Variable focus, zoom
- Image intensifiers
- Polarization camera's

Camera calibration and fusion

Image processing and camera simulation



Machine learning

Camera simulation

1. The problem with vanilla GPs

The problem with vanilla GPs

The posterior is a GP!!!

The problem with vanilla GPs

The posterior is a GP!!!

Problem locally: A prediction at $\mathbf{x}_*$ is a Gaussian distribution, ranging from $-\infty$ to $+\infty$. What about velocities, weights, heights, concentrations in %, ... ?

The problem with vanilla GPs

The posterior is a GP!!!

Problem locally: A prediction at $\mathbf{x}_*$ is a Gaussian distribution, ranging from $-\infty$ to $+\infty$. What about velocities, weights, heights, concentrations in %, ... ?

Problem globally: Kernel determines shape of the functions. Samples from posterior don't automatically obey monotonicity, convexity, boundary conditions, differential equations (heat, forces, ...), ...

The problem with vanilla GPs

The posterior is a GP!!!

Problem locally: A prediction at $\mathbf{x}_*$ is a Gaussian distribution, ranging from $-\infty$ to $+\infty$. What about velocities, weights, heights, concentrations in %, ... ?

Problem globally: Kernel determines shape of the functions. Samples from posterior don't automatically obey monotonicity, convexity, boundary conditions, differential equations (heat, forces, ...), ...

Problem multi-output: Relationships between outputs are not built-in. What about zero curl or divergence in a vector field? What about unit norm vectors?

The problem with vanilla GPs

The posterior is a GP!!!

Problem locally: A prediction at $\mathbf{x}_*$ is a Gaussian distribution, ranging from $-\infty$ to $+\infty$. What about velocities, weights, heights, concentrations in %, ... ?

Problem globally: Kernel determines shape of the functions. Samples from posterior don't automatically obey monotonicity, convexity, boundary conditions, differential equations (heat, forces, ...), ...

Problem multi-output: Relationships between outputs are not built-in. What about zero curl or divergence in a vector field? What about unit norm vectors?

What about your data? **Does it matter**?

2. Problem Locally - Bound Constraints

Bound Constraints

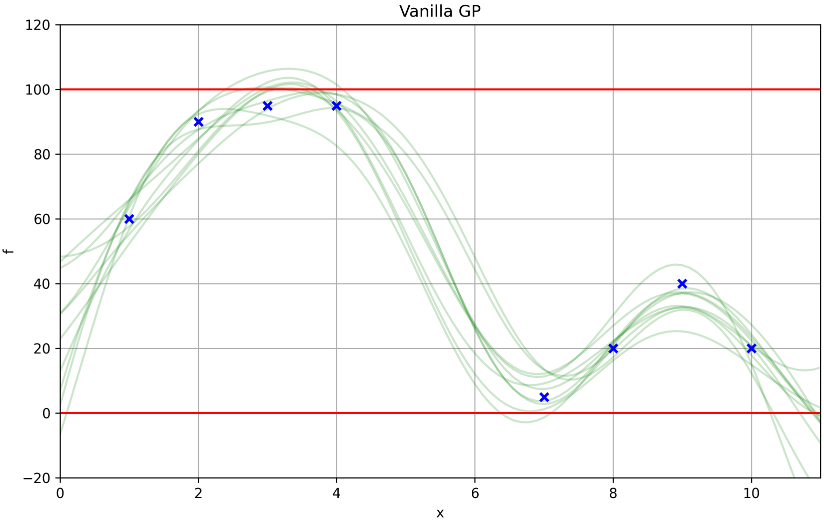
$$a \leq f(x) \leq b$$

Bound Constraints

$$0 \leq f(x) \leq 100$$

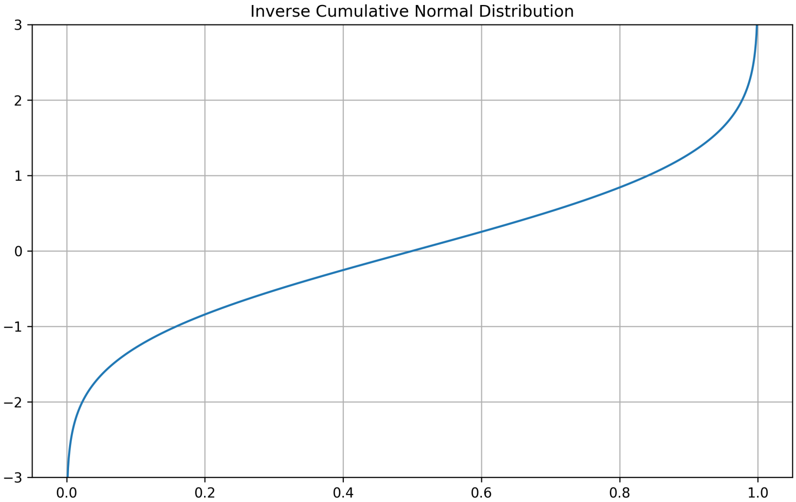
Bound Constraints

$$0 \leq f(x) \leq 100$$



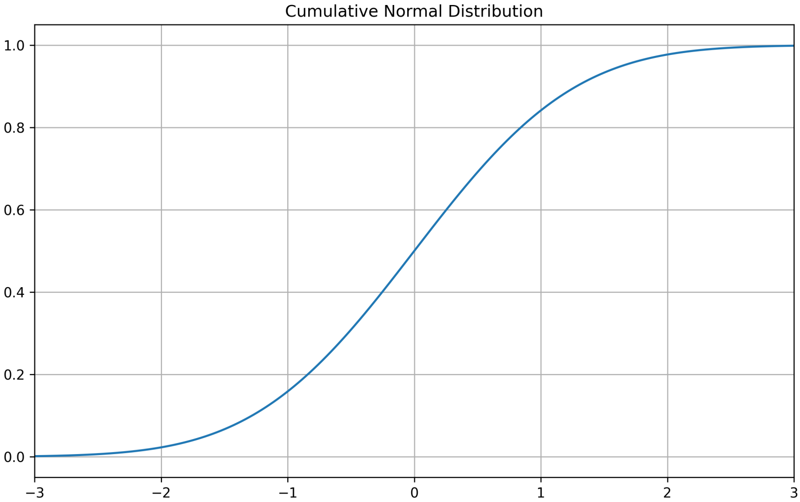
Bound Constraints - Warping

$$0 \leq f(x) \leq 100 \ , \ u_i = \Phi^{-1}(y_i), \ [0,1] \rightarrow [-\infty, +\infty]$$



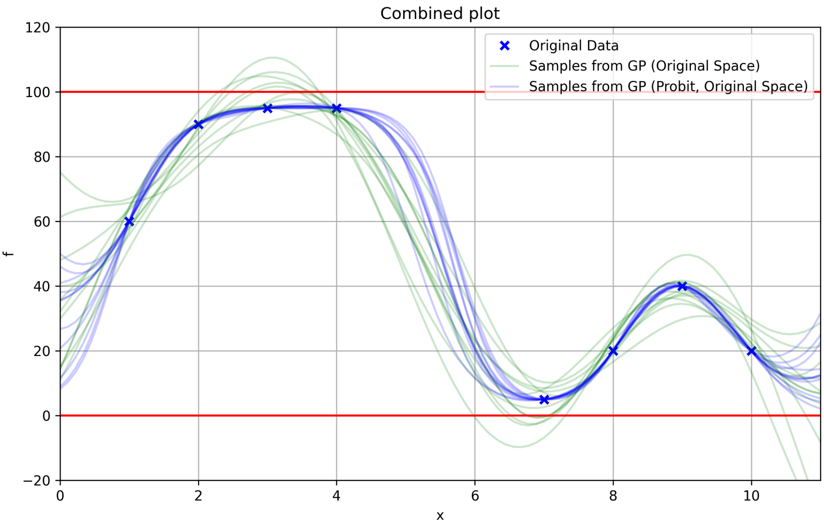
Bound Constraints - Warping

train GP on $\{x_i, u_i\}$, warp prediction back via Φ



Bound Constraints - Warping

$$0 \leq f(x) \leq 100$$



University of Antwerp
Faculty of Applied Engineering

Bound Constraints - Likelihood Tinkering

$$p(\mathbf{f}|X, y, \theta) = \frac{p(y|X,\mathbf{f},\theta)p(\mathbf{f}|X,\theta)}{p(y|X,\theta)}$$

3. Problem Globally - Differential Equation Constraints

Differential Equation Constraints with Data

$\mathcal{L}u = f$, with data on u and f, and $\mathcal{L}$ linear partial differential operator

$$u \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$$

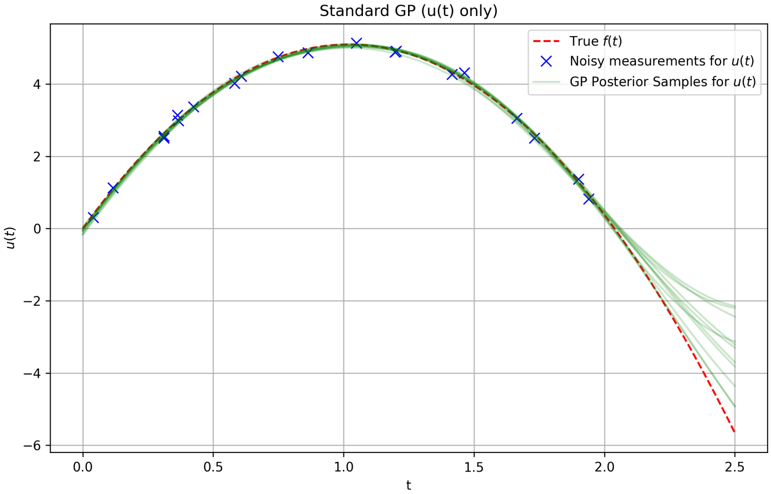
$$\mathcal{L}u \sim \mathcal{GP}(\mathcal{L}_{\mathbf{x}} m(\mathbf{x}), \mathcal{L}_{\mathbf{x}}\mathcal{L}_{\mathbf{x}'} k(\mathbf{x}, \mathbf{x}'))$$

$$\begin{bmatrix} u(X_1) \\ f(X_2) \end{bmatrix} \sim \mathcal{GP}\left(\begin{bmatrix} m(X_1) \\ \mathcal{L}m(X_2) \end{bmatrix}, \begin{bmatrix} K_{11}(X_1, X_1) & K_{12}(X_1, X_2) \\ K_{21}(X_2, X_1) & K_{22}(X_2, X_2) \end{bmatrix} \right)$$

$$k\left(\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{x}_1' \\ \mathbf{x}_2' \end{bmatrix} \right) = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1') & \mathcal{L}_{\mathbf{x}'} k(\mathbf{x}_1, \mathbf{x}_2') \\ \mathcal{L}_{\mathbf{x}} k(\mathbf{x}_2, \mathbf{x}_1') & \mathcal{L}_{\mathbf{x}}\mathcal{L}_{\mathbf{x}'} k(\mathbf{x}_2, \mathbf{x}_2') \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$$

Differential Equation Constraints with Data

Kick a ball: height $u(t) = at - \frac{g}{2}t^2$ for some unknown a.
We observe u_i.

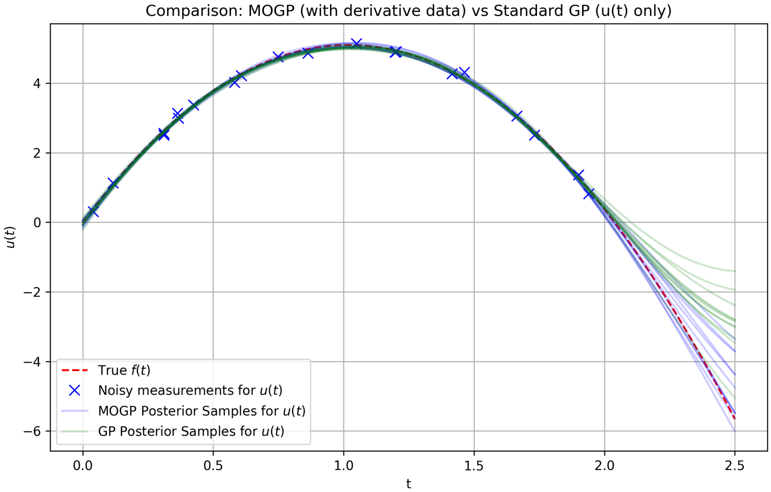


University of Antwerp | Faculty of Applied Engineering

Differential Equation Constraints with Data

Kick a ball: height $u(t) = at - \frac{g}{2}t^2$ for some unknown a.

We observe u_i and know $f(t) = \frac{d^2u}{dt^2} = -g = f_i$.

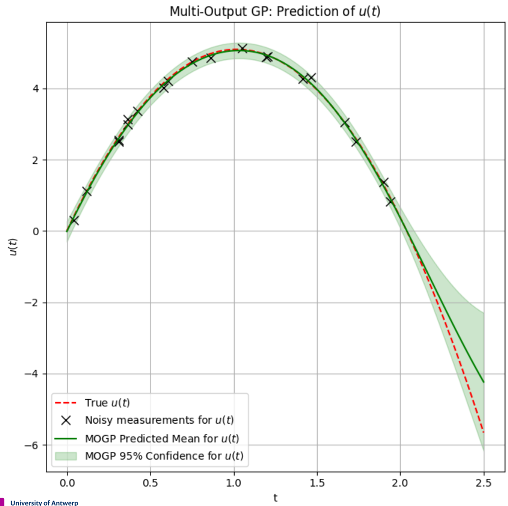


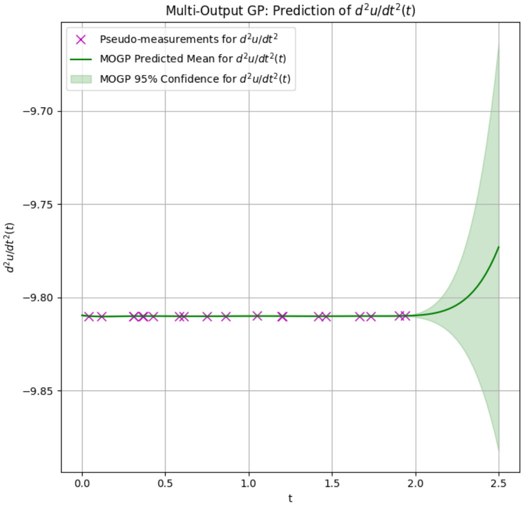
University of Antwerp | Faculty of Applied Engineering

Differential Equation Constraints with Data

Kick a ball: height $u(t) = at - \frac{g}{2}t^2$ for some unknown a.

We observe u_i and know $f(t) = \frac{d^2u}{dt^2} = -g = f_i$.





University of Antwerp
Faculty of Applied Engineering

Monotonicity Constraint

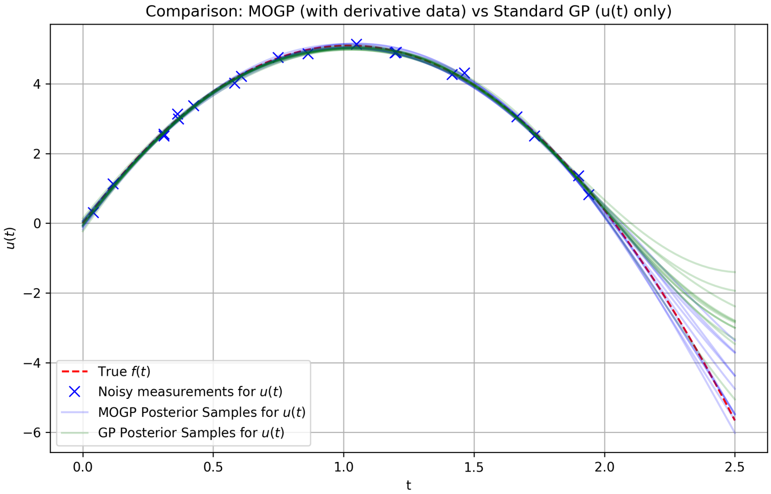
$$\frac{df}{dt} > 0$$

Convexity Constraint

$$\frac{d^2 f}{dt^2} > 0$$

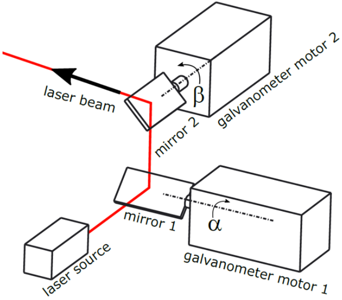
Concavity Constraint

$$\frac{d^2 f}{dt^2} < 0$$

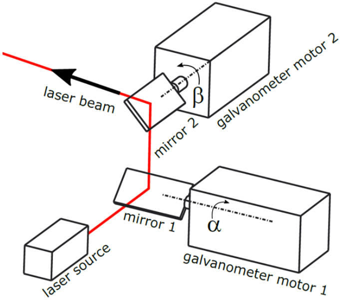


4. Problem Multi-Output - e.g. Predict Lines in 3D

Galvanometric Laser Scanner

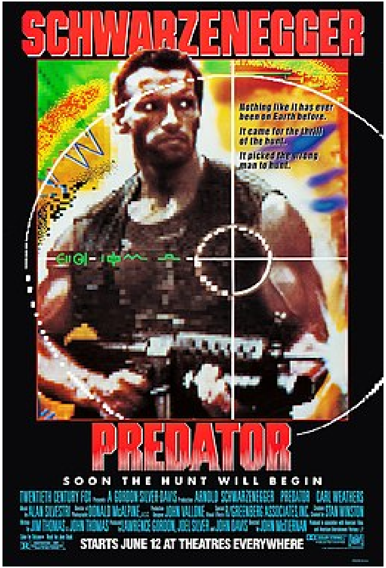


Galvanometric Laser Scanner

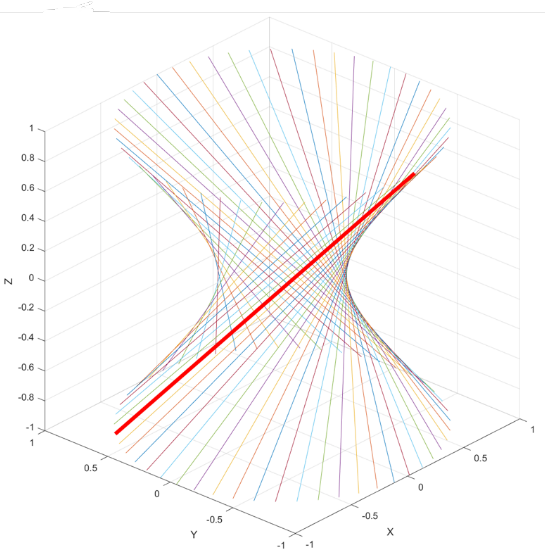


$$(\alpha, \beta) \rightarrow L$$

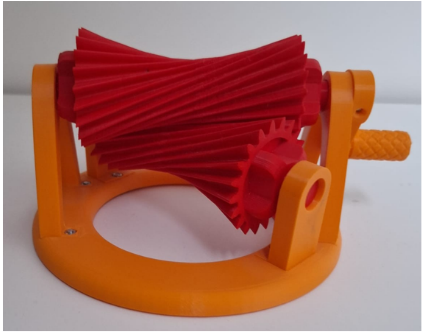
Predator


SCHWARZENEGGER
Nothing like it has ever been on Earth before.
It came for the thrill of the hunt.
It picked the wrong man to hunt.
PREDATOR
SOON THE HUNT WILL BEGIN
STARTS JUNE 12 AT THEATRES EVERYWHERE

Hyperboloid



Hyperboloids

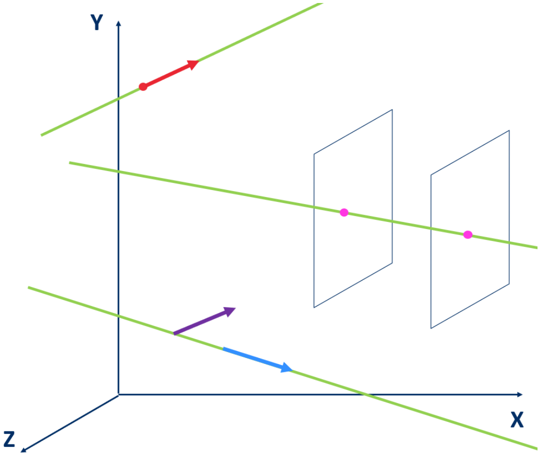


Hyperboloid Bridge

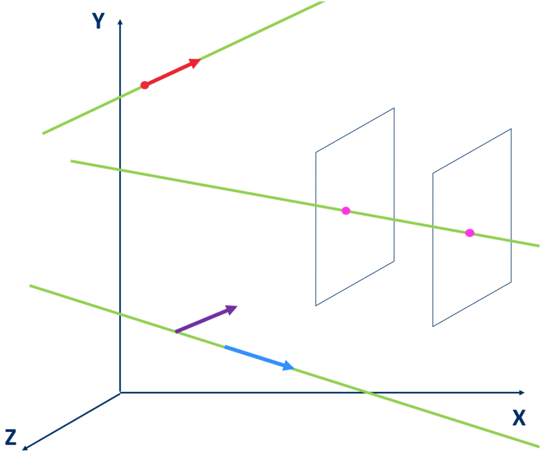
Corporation Street Bridge

Coordinates	53°29′01″N 2°14′36″W
Crosses	Corporation Street
Locale	Manchester, England

Plücker Coordinates



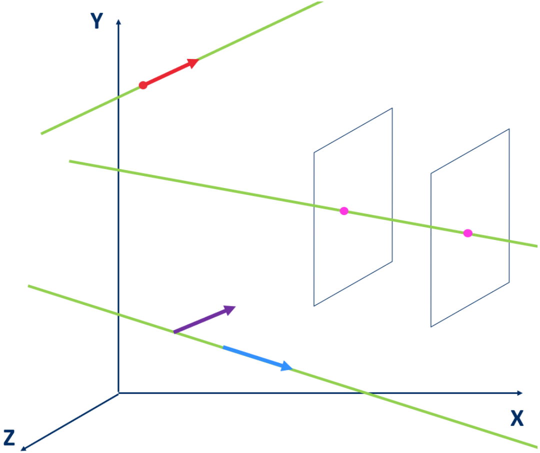
Plücker Coordinates



L given by moment $\mathbf{m}$ and direction $\mathbf{d}$: $(l_1 : l_2 : l_3 : l_4 : l_5 : l_6) \in \mathbb{P}^5$

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$ (Grassmann-Plücker relationship)

Plücker Coordinates



L given by moment $\mathbf{m}$ and direction $\mathbf{d}$: $(l_1 : l_2 : l_3 : l_4 : l_5 : l_6) \in \mathbb{P}^5$

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$ (Grassmann-Plücker relationship)

$L \in M_2^4$ (Klein Quadric) $\subset \mathbb{P}^5$

Parallel GPs screw this up

$$(\alpha, \beta) \stackrel{GP_1}{\rightarrow} l_1$$
$$(\alpha, \beta) \stackrel{GP_2}{\rightarrow} l_2$$
$$(\alpha, \beta) \stackrel{GP_3}{\rightarrow} l_3$$
$$(\alpha, \beta) \stackrel{GP_4}{\rightarrow} l_4$$
$$(\alpha, \beta) \stackrel{GP_5}{\rightarrow} l_5$$
$$(\alpha, \beta) \stackrel{GP_6}{\rightarrow} l_6$$

Parallel GPs screw this up

$$(\alpha, \beta) \stackrel{GP_1}{\rightarrow} l_1$$

$$(\alpha, \beta) \stackrel{GP_2}{\rightarrow} l_2$$

$$(\alpha, \beta) \stackrel{GP_3}{\rightarrow} l_3$$

$$(\alpha, \beta) \stackrel{GP_4}{\rightarrow} l_4$$

$$(\alpha, \beta) \stackrel{GP_5}{\rightarrow} l_5$$

$$(\alpha, \beta) \stackrel{GP_6}{\rightarrow} l_6$$

$$l_1 l_4 + l_2 l_5 + l_3 l_6 = ?$$

Parallel GPs screw this up

$$(\alpha, \beta) \stackrel{GP_1}{\rightarrow} l_1$$

$$(\alpha, \beta) \stackrel{GP_2}{\rightarrow} l_2$$

$$(\alpha, \beta) \stackrel{GP_3}{\rightarrow} l_3$$

$$(\alpha, \beta) \stackrel{GP_4}{\rightarrow} l_4$$

$$(\alpha, \beta) \stackrel{GP_5}{\rightarrow} l_5$$

$$(\alpha, \beta) \stackrel{GP_6}{\rightarrow} l_6$$

$$l_1 l_4 + l_2 l_5 + l_3 l_6 = ?$$

$$l_1 l_4 + l_2 l_5 + l_3 l_6 \neq 0 \Rightarrow (l_1 : l_2 : l_3 : l_4 : l_5 : l_6) \text{ is a screw, } \notin M_2^4$$

NOT a line in 3D!!!

Constraints on predictions for L

$$l_1 l_1 + l_2 l_2 + l_3 l_3 = 1$$

$$l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$$

5. Unit Norm: $l_1 l_1 + l_2 l_2 + l_3 l_3 = 1$

Input (α, β)

Periodic kernel in general:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{2}{\ell^2}\sin^2\left(\pi\frac{|\mathbf{x}-\mathbf{x}'|}{p}\right)\right)$$

Periodic kernel for $\mathbf{x} = (\alpha, \beta)$, ARD and period 2π:

$$k_{PER}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{2}{l_\alpha^2}\sin^2\left(\frac{|\alpha-\alpha'|}{2}\right)\right)\exp\left(-\frac{2}{l_\beta^2}\sin^2\left(\frac{|\beta-\beta'|}{2}\right)\right)$$

Unit Norm for Direction Vector $\mathbf{d}$

$(\alpha, \beta) \rightarrow \mathbf{d} \in \mathbb{R}^3$

Unit Norm for Direction Vector $\mathbf{d}$

$(\alpha, \beta) \to \mathbf{d} \in \mathbb{R}^3$, constraint: $\|\mathbf{d}\| = 1, \mathbf{d} \in \mathbb{S}^2$

Unit Norm for Direction Vector $\mathbf{d}$

$(\alpha, \beta) \to \mathbf{d} \in \mathbb{R}^3$, constraint: $\|\mathbf{d}\| = 1, \mathbf{d} \in \mathbb{S}^2$

New kernel for $\mathbf{x} = (\alpha, \beta, \text{dim})$, with dim $\in 1, 2, 3$:

$$k_{PER3D}(\mathbf{x}, \mathbf{x}') = \begin{cases} k_{PER}([\alpha, \beta], [\alpha', \beta']), & \text{for dim} = \text{dim}' \\ 0 & \text{for dim} \neq \text{dim}' \end{cases}$$

Unit Norm for Direction Vector $\mathbf{d}$

$(\alpha, \beta) \to \mathbf{d} \in \mathbb{R}^3$, constraint: $\|\mathbf{d}\| = 1, \mathbf{d} \in \mathbb{S}^2$

New kernel for $\mathbf{x} = (\alpha, \beta, \text{dim})$, with dim $\in 1, 2, 3$:

$$k_{PER3D}(\mathbf{x}, \mathbf{x}') = \begin{cases} k_{PER}([\alpha, \beta], [\alpha', \beta']), & \text{for dim} = \text{dim}' \\ 0 & \text{for dim} \neq \text{dim}' \end{cases}$$

$$f_1 : x = \cos(\alpha)\sin(\beta),$$
$$f_2 : y = \sin(\alpha)\sin(\beta),$$
$$f_3 : z = \cos(\beta)$$

Unit Norm for Direction Vector $\mathbf{d}$

$(\alpha, \beta) \rightarrow \mathbf{d} \in \mathbb{R}^3$, constraint: $\|\mathbf{d}\| = 1, \mathbf{d} \in \mathbb{S}^2$

New kernel for $\mathbf{x} = (\alpha, \beta, \text{dim})$, with $\text{dim} \in 1, 2, 3$:

$$k_{PER3D}(\mathbf{x}, \mathbf{x}') = \begin{cases} k_{PER}([\alpha, \beta], [\alpha', \beta']), & \text{for dim} = \text{dim}' \\ 0 & \text{for dim} \neq \text{dim}' \end{cases}$$

$$\begin{aligned} f_1 &: x = \cos(\alpha)\sin(\beta), \\ f_2 &: y = \sin(\alpha)\sin(\beta), \\ f_3 &: z = \cos(\beta) \end{aligned}$$

$$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$$

Unit Norm for Direction Vector d

$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$

Unit Norm for Direction Vector d

$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$, $\mathcal{F}_{\mathbf{x}}[\mathbf{f}(\mathbf{x})] = \mathbf{0}$

Unit Norm for Direction Vector $\mathbf{d}$

$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$, $\mathcal{F}_{\mathbf{x}}[\mathbf{f}(\mathbf{x})] = \mathbf{0}$, $\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})] = \mathbf{f}(\mathbf{x})$

Unit Norm for Direction Vector $\mathbf{d}$

$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$, $\mathcal{F}_{\mathbf{x}}[\mathbf{f}(\mathbf{x})] = \mathbf{0}$, $\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})] = \mathbf{f}(\mathbf{x})$, $\mathcal{F}_{\mathbf{x}}[\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})]] = \mathbf{0}$,

Unit Norm for Direction Vector d

$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$, $\mathcal{F}_{\mathbf{x}}[\mathbf{f}(\mathbf{x})] = \mathbf{0}$, $\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})] = \mathbf{f}(\mathbf{x})$, $\mathcal{F}_{\mathbf{x}}[\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})]] = \mathbf{0}$,

$$\mathbf{g}(\mathbf{x}) \sim \mathcal{GP}(\mu_{\mathbf{g}}, k_{\mathbf{g}}), \quad \mathbf{f}(\mathbf{x}) = \mathcal{G}_{\mathbf{x}}\mathbf{g}(\mathbf{x}) \sim \mathcal{GP}(\mathcal{G}_{\mathbf{x}}\mu_{\mathbf{g}}, \mathcal{G}_{\mathbf{x}} k_{\mathbf{g}} \mathcal{G}_{\mathbf{x}'}^T)$$

Unit Norm for Direction Vector $\mathbf{d}$

$f_1 - \frac{\partial f_2}{\partial \alpha} + \frac{\partial f_3}{\partial \alpha} = 0$, $\mathcal{F}_{\mathbf{x}}[\mathbf{f}(\mathbf{x})] = \mathbf{0}$, $\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})] = \mathbf{f}(\mathbf{x})$, $\mathcal{F}_{\mathbf{x}}[\mathcal{G}_{\mathbf{x}}[\mathbf{g}(\mathbf{x})]] = \mathbf{0}$,

$$\mathbf{g}(\mathbf{x}) \sim \mathcal{GP}(\mu_{\mathbf{g}}, k_{\mathbf{g}}), \quad \mathbf{f}(\mathbf{x}) = \mathcal{G}_{\mathbf{x}}\mathbf{g}(\mathbf{x}) \sim \mathcal{GP}(\mathcal{G}_{\mathbf{x}}\mu_{\mathbf{g}}, \mathcal{G}_{\mathbf{x}} k_{\mathbf{g}} \mathcal{G}_{\mathbf{x}'}^T)$$

$$\mathcal{F}_{\mathbf{x}} = \left[1, -\frac{\partial}{\partial_\alpha}, \frac{\partial}{\partial_\alpha}\right]$$

$$\mathcal{G}_{\mathbf{x}} = \left[\frac{\partial}{\partial_\alpha}, 1, 0\right]^T, \quad \mathcal{F}_{\mathbf{x}}\mathcal{G}_{\mathbf{x}} = \mathbf{0}$$

$$\mathcal{G}_{\mathbf{x}}\mathcal{G}_{\mathbf{x}'}^T = \begin{bmatrix} \frac{\partial^2}{\partial_\alpha \partial_{\alpha'}} & \frac{\partial}{\partial_\alpha} & 0 \\ \frac{\partial}{\partial_{\alpha'}} & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Unit Norm for Direction Vector d

$$\mathbf{x} = (\alpha, \beta, \text{dim}), \text{ with dim} \in 1, 2, 3$$

$$k_{PER3D}(\mathbf{x}, \mathbf{x}') = \begin{cases} k_{PER}([\alpha, \beta], [\alpha', \beta']), & \text{for dim} = \text{dim}' \\ 0 & \text{for dim} \neq \text{dim}' \end{cases}$$

University of Antwerp | Faculty of Applied Engineering

Unit Norm for Direction Vector d

$$\mathbf{x} = (\alpha, \beta, \text{dim}), \text{ with dim} \in 1, 2, 3$$

$$k_{PER3D}(\mathbf{x}, \mathbf{x}') = \begin{cases} k_{PER}([\alpha, \beta], [\alpha', \beta']), & \text{for dim} = \text{dim}' \\ 0 & \text{for dim} \neq \text{dim}' \end{cases}$$

$$k_{COM}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{2}{l_\alpha^2} \sin^2\left(\frac{\alpha - \alpha'}{2}\right)\right) \cdot \exp\left(-\frac{2}{l_\beta^2} \sin^2\left(\frac{\beta - \beta'}{2}\right)\right) \cdot \exp\left(-\frac{(\text{dim} - \text{dim}')^2}{\epsilon^2}\right)$$

Unit Norm for Direction Vector d

$$k_{DIF} = \mathcal{G}_{\mathbf{x}} k_{COM} \mathcal{G}_{\mathbf{x}'}^{T} = \begin{bmatrix} A & B & 0 \\ C & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} k_{COM},$$

in which, with $\gamma = \frac{\alpha - \alpha'}{2}$ (for brevity in notation),

$$A = -\frac{\sin^2(\gamma)\cos^2(\gamma)}{l_\alpha^4} + \frac{\cos^2(\gamma)}{l_\alpha^2} - \frac{\sin^2(\gamma)}{l_\alpha^2},$$

$$B = -\frac{2}{l_\alpha^2}\sin(\gamma)\cos(\gamma),$$

$$C = \frac{2}{l_\alpha^2}\sin(\gamma)\cos(\gamma)$$

Unit Norm for Direction Vector $\mathbf{d}$

$$k_{DIF} = \mathcal{G}_{\mathbf{x}} k_{COM} \mathcal{G}_{\mathbf{x}'}^{T} = \begin{bmatrix} A & B & 0 \\ C & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} k_{COM},$$

in which, with $\gamma = \frac{\alpha - \alpha'}{2}$ (for brevity in notation),

$$A = -\frac{\sin^2(\gamma)\cos^2(\gamma)}{l_\alpha^4} + \frac{\cos^2(\gamma)}{l_\alpha^2} - \frac{\sin^2(\gamma)}{l_\alpha^2},$$

$$B = -\frac{2}{l_\alpha^2}\sin(\gamma)\cos(\gamma),$$

$$C = \frac{2}{l_\alpha^2}\sin(\gamma)\cos(\gamma)$$

$$(\alpha, \beta) \rightarrow \mathbf{d} \in \mathbb{S}^2$$

$$\mathbf{f}(\mathbf{x}) = \mathcal{G}_{\mathbf{x}}\mathbf{g}(\mathbf{x}) \sim \mathcal{GP}(\mathcal{G}_{\mathbf{x}}\mu_{\mathbf{g}}, \mathcal{G}_{\mathbf{x}} k_{\mathbf{g}} \mathcal{G}_{\mathbf{x}'}^{T})$$

6. Grassmann-Plücker Relationship: $l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$

Linear Constraint in general

A set of training examples $\{y_1, ..., y_n\}$ that satisfy the linear constraints $\mathbf{A}y_i = \mathbf{b}$ will result in a GP for which the mean prediction $\mu(\mathbf{x}_*)$ also satisfies $\mathbf{A}\mu(\mathbf{x}_*) = \mathbf{b}$.

Linear Constraint in general

A set of training examples $\{y_1, ..., y_n\}$ that satisfy the linear constraints $\mathbf{A}y_i = \mathbf{b}$ will result in a GP for which the mean prediction $\mu(\mathbf{x}_*)$ also satisfies $\mathbf{A}\mu(\mathbf{x}_*) = \mathbf{b}$.

$$\mu(\mathbf{x}_*) \quad = \quad \mathbf{m}_0(\mathbf{x}_*) + \mathbf{K}(\mathbf{x}_*, \mathbf{X}) \left[\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}\right]^{-1} (\mathbf{y} - \mathbf{m}(\mathbf{x}))$$

Quadratic Constraint

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$, quadratic in l_i, but linear in $l_i l_j$

Quadratic Constraint

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$, quadratic in l_i, but linear in $l_i l_j$

Let $\mathbf{z} = (l_1, l_2, l_3, l_4, l_5, l_6)$ and $\mathbf{Q} = \mathbf{z}^T \mathbf{z} =$

$$\begin{bmatrix} l_1 l_1 & l_1 l_2 & l_1 l_3 & l_1 l_4 & l_1 l_5 & l_1 l_6 \\ l_2 l_1 & l_2 l_2 & l_2 l_3 & l_2 l_4 & l_2 l_5 & l_2 l_6 \\ l_3 l_1 & l_3 l_2 & l_3 l_3 & l_3 l_4 & l_3 l_5 & l_3 l_6 \\ l_4 l_1 & l_4 l_2 & l_4 l_3 & l_4 l_4 & l_4 l_5 & l_4 l_6 \\ l_5 l_1 & l_5 l_2 & l_5 l_3 & l_5 l_4 & l_5 l_5 & l_5 l_6 \\ l_6 l_1 & l_6 l_2 & l_6 l_3 & l_6 l_4 & l_6 l_5 & l_6 l_6 \end{bmatrix}$$

Quadratic Constraint

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$, quadratic in l_i, but linear in $l_i l_j$

$$\text{Let } \mathbf{z} = (l_1, l_2, l_3, l_4, l_5, l_6) \text{ and } \mathbf{Q} = \mathbf{z}^T \mathbf{z} = \begin{bmatrix} l_1 l_1 & l_1 l_2 & l_1 l_3 & l_1 l_4 & l_1 l_5 & l_1 l_6 \\ l_2 l_1 & l_2 l_2 & l_2 l_3 & l_2 l_4 & l_2 l_5 & l_2 l_6 \\ l_3 l_1 & l_3 l_2 & l_3 l_3 & l_3 l_4 & l_3 l_5 & l_3 l_6 \\ l_4 l_1 & l_4 l_2 & l_4 l_3 & l_4 l_4 & l_4 l_5 & l_4 l_6 \\ l_5 l_1 & l_5 l_2 & l_5 l_3 & l_5 l_4 & l_5 l_5 & l_5 l_6 \\ l_6 l_1 & l_6 l_2 & l_6 l_3 & l_6 l_4 & l_6 l_5 & l_6 l_6 \end{bmatrix}$$

Veronese mapping, $V_{52} : \mathbb{P}^5 \to \mathbb{P}^{20}$, $V_{52}(\mathbb{P}^5) \subset \mathbb{P}^{20}$

$\mathbf{Q}$ is rank 1, $\det(\mathbf{Q}) = 0$, all 2×2 minors $= 0$, 210 (independent) quadratic equations

Quadratic Constraint

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$, quadratic in l_i, but linear in $l_i l_j$

Let $\mathbf{z} = (l_1, l_2, l_3, l_4, l_5, l_6)$ and $\mathbf{Q} = \mathbf{z}^T \mathbf{z} = \begin{bmatrix} l_1 l_1 & l_1 l_2 & l_1 l_3 & l_1 l_4 & l_1 l_5 & l_1 l_6 \\ l_2 l_1 & l_2 l_2 & l_2 l_3 & l_2 l_4 & l_2 l_5 & l_2 l_6 \\ l_3 l_1 & l_3 l_2 & l_3 l_3 & l_3 l_4 & l_3 l_5 & l_3 l_6 \\ l_4 l_1 & l_4 l_2 & l_4 l_3 & l_4 l_4 & l_4 l_5 & l_4 l_6 \\ l_5 l_1 & l_5 l_2 & l_5 l_3 & l_5 l_4 & l_5 l_5 & l_5 l_6 \\ l_6 l_1 & l_6 l_2 & l_6 l_3 & l_6 l_4 & l_6 l_5 & l_6 l_6 \end{bmatrix}$

Veronese mapping, $V_{52} : \mathbb{P}^5 \rightarrow \mathbb{P}^{20}$, $V_{52}(\mathbb{P}^5) \subset \mathbb{P}^{20}$
$\mathbf{Q}$ is rank 1, $\det(\mathbf{Q}) = 0$, all 2×2 minors $= 0$, 210 (independent) quadratic equations

Formulate $\mathbf{y} \in \mathbb{R}^{21}$ as the concatenation of **upper triangular elements**

$$\mathbf{y} = [\mathbf{Q}_{11}, ..., \mathbf{Q}_{ij}, ..., \mathbf{Q}_{66}]^T, \text{with } i \leq j$$

Quadratic Constraint

$\mathbf{m} \perp \mathbf{d} \Rightarrow l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$, quadratic in l_i, but linear in $l_i l_j$

Let $\mathbf{z} = (l_1, l_2, l_3, l_4, l_5, l_6)$ and $\mathbf{Q} = \mathbf{z}^T \mathbf{z} = \begin{bmatrix} l_1 l_1 & l_1 l_2 & l_1 l_3 & l_1 l_4 & l_1 l_5 & l_1 l_6 \\ l_2 l_1 & l_2 l_2 & l_2 l_3 & l_2 l_4 & l_2 l_5 & l_2 l_6 \\ l_3 l_1 & l_3 l_2 & l_3 l_3 & l_3 l_4 & l_3 l_5 & l_3 l_6 \\ l_4 l_1 & l_4 l_2 & l_4 l_3 & l_4 l_4 & l_4 l_5 & l_4 l_6 \\ l_5 l_1 & l_5 l_2 & l_5 l_3 & l_5 l_4 & l_5 l_5 & l_5 l_6 \\ l_6 l_1 & l_6 l_2 & l_6 l_3 & l_6 l_4 & l_6 l_5 & l_6 l_6 \end{bmatrix}$

Veronese mapping, $V_{52} : \mathbb{P}^5 \to \mathbb{P}^{20}$, $V_{52}(\mathbb{P}^5) \subset \mathbb{P}^{20}$
$\mathbf{Q}$ is rank 1, $\det(\mathbf{Q}) = 0$, all 2×2 minors $= 0$, 210 (independent) quadratic equations

Formulate $\mathbf{y} \in \mathbb{R}^{21}$ as the concatenation of **upper triangular elements**

$$\mathbf{y} = [\mathbf{Q}_{11}, \ldots, \mathbf{Q}_{ij}, \ldots, \mathbf{Q}_{66}]^T, \text{with } i \leq j$$

$\mathbf{x} = (\alpha, \beta, \dim)$, with $\dim \in \{1, 2, \ldots, 21\}$

Quadratic Constraint

$\mathbf{A}\mathbf{y}_i = \mathbf{b}$, two constraints: $l_1 l_1 + l_2 l_2 + l_3 l_3 = 1, l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$

Quadratic Constraint

$\mathbf{A}\mathbf{y}_i = \mathbf{b}$, two constraints: $l_1 l_1 + l_2 l_2 + l_3 l_3 = 1, l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \end{bmatrix}_{2\times 21} \begin{bmatrix} l_1 l_1 \\ l_1 l_2 \\ l_1 l_3 \\ l_1 l_4 \\ l_1 l_5 \\ l_1 l_6 \\ l_2 l_2 \\ l_2 l_1 \\ \vdots \\ l_5 l_5 \\ l_5 l_6 \\ l_6 l_6 \end{bmatrix}_{21\times 1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Quadratic Constraint

$\mathbf{A}\mathbf{y}_i = \mathbf{b}$, two constraints: $l_1 l_1 + l_2 l_2 + l_3 l_3 = 1, l_1 l_4 + l_2 l_5 + l_3 l_6 = 0$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \end{bmatrix}_{2\times 21} \begin{bmatrix} l_1 l_1 \\ l_1 l_2 \\ l_1 l_3 \\ l_1 l_4 \\ l_1 l_5 \\ l_1 l_6 \\ l_2 l_2 \\ l_2 l_1 \\ \vdots \\ l_5 l_5 \\ l_5 l_6 \\ l_6 l_6 \end{bmatrix}_{21\times 1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

For $\mu(\mathbf{x}_*)$ satisfies $\mathbf{A}\mu(\mathbf{x}_*) = \mathbf{b}$

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$
$\exists \mathbf{y}_* \notin V_{52}(\mathbb{P}^5)$, let alone corresponding to point a on M_2^4 in $\mathbb{P}^5$!

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$
$\exists \mathbf{y}_* \notin V_{52}(\mathbb{P}^5)$, let alone corresponding to point a on M_2^4 in $\mathbb{P}^5$!
E.g.: $(1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 666 \quad 0)$

University of Antwerp | Faculty of Applied Engineering

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$
$\exists \mathbf{y}_* \notin V_{52}(\mathbb{P}^5)$, let alone corresponding to point a on M_2^4 in $\mathbb{P}^5$!
E.g.: $(1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 666 \quad 0)$
$\mathbf{Q} = \mathbf{z}^T\mathbf{z}$ needs to be rank 1.

University of Antwerp | Faculty of Applied Engineering

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$
$\exists \mathbf{y}_* \notin V_{52}(\mathbb{P}^5)$, let alone corresponding to point a on M_2^4 in $\mathbb{P}^5$!
E.g.: $(1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 666 \quad 0)$
$\mathbf{Q} = \mathbf{z}^T\mathbf{z}$ needs to be rank 1.

$$\mathbf{Q}_* = \begin{bmatrix} y_{*1} & y_{*2} & \cdots & y_{*6} \\ y_{*2} & y_{*7} & \cdots & y_{*11} \\ \vdots & \vdots & \vdots & \vdots \\ y_{*6} & y_{*11} & \cdots & y_{*21} \end{bmatrix} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$
$\exists \mathbf{y}_* \notin V_{52}(\mathbb{P}^5)$, let alone corresponding to point a on M_2^4 in $\mathbb{P}^5$!
E.g.: $(1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 666 \quad 0)$
$\mathbf{Q} = \mathbf{z}^T\mathbf{z}$ needs to be rank 1.

$$\mathbf{Q}_* = \begin{bmatrix} y_{*1} & y_{*2} & \cdots & y_{*6} \\ y_{*2} & y_{*7} & \cdots & y_{*11} \\ \vdots & \vdots & \vdots & \vdots \\ y_{*6} & y_{*11} & \cdots & y_{*21} \end{bmatrix} = \mathbf{U}\Sigma\mathbf{V}^T$$

$$\mathbf{z}_* = \sqrt{\Sigma_{(1,1)}}\mathbf{V}_{(:,1)}$$

University of Antwerp
Faculty of Applied Engineering

Quadratic Constraint

We don't want the 21D $\mu(\mathbf{x}_*) = \mathbf{y}_*$, but the 6D $\mathbf{z}_*$.

Problem: $\mathbf{A}\mathbf{y}_* = \mathbf{b} \iff \mathbf{y}_* \in H1 \cap H2$ in $\mathbb{R}^{21}$ (one hyperplane for every row in $\mathbf{A}$)
$\dim(H1 \cap H2) = 21 - 1 - 1 = 19$, $\dim(V_{52}(\mathbb{P}^5)) = 5$
$\exists \mathbf{y}_* \notin V_{52}(\mathbb{P}^5)$, let alone corresponding to point a on M_2^4 in $\mathbb{P}^5$!
E.g.: $(1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 666 \quad 0)$
$\mathbf{Q} = \mathbf{z}^T\mathbf{z}$ needs to be rank 1.

$$\mathbf{Q}_* = \begin{bmatrix} y_{*1} & y_{*2} & \cdots & y_{*6} \\ y_{*2} & y_{*7} & \cdots & y_{*11} \\ \vdots & \vdots & \vdots & \vdots \\ y_{*6} & y_{*11} & \cdots & y_{*21} \end{bmatrix} = \mathbf{U}\Sigma\mathbf{V}^T$$

$$\mathbf{z}_* = \sqrt{\Sigma_{(1,1)}}\mathbf{V}_{(:,1)}$$

Sign ambiguity can be solved from context.

University of Antwerp | Faculty of Applied Engineering

No free lunch

$$\mathcal{O}(n^3)$$

No free lunch

$$\mathcal{O}(n^3)$$

$$n = 81 \Rightarrow n \times 21 = 1701$$

No free lunch

$$\mathcal{O}(n^3)$$

$$n = 81 \Rightarrow n \times 21 = 1701$$

6 $\times$ 2D GPs, 6 predictions per line

$$\Rightarrow$$

1 $\times$ 3D GP, 21 predictions per line plus SVD

The problem with vanilla GPs

The posterior is a GP!!!

Problem locally: A prediction at $\mathbf{x}_*$ is a Gaussian distribution, ranging from $-\infty$ to $+\infty$. What about velocities, weights, heights, concentrations in %, ... ?

Problem globally: Kernel determines shape of the functions. Samples from posterior don't automatically obey monotonicity, convexity, boundary conditions, differential equations (heat, forces, ...), ...

Problem multi-output: Relationships between outputs are not built-in. What about zero curl or divergence in a vector field? What about unit norm vectors?

What about your data? **Does it matter**?

Further reading

Swiler, L. P., Gulian, M., Frankel, A. L., Safta, C., & Jakeman, J. D. (2020). **A survey of constrained Gaussian process regression: Approaches and implementation challenges.** Journal of Machine Learning for Modeling and Computing, 1(2).

Salzmann, M., & Urtasun, R. (2010). **Implicitly constrained Gaussian process regression for monocular non-rigid pose estimation.** Advances in neural information processing systems, 23.

De Boi, I., Sels, S., De Moor, O., Vanlanduit, S., & Penne, R. (2022). **Input and output manifold constrained Gaussian process regression for galvanometric setup calibration.** IEEE Transactions on Instrumentation and Measurement, 71, 1-8.

InViLab

ivan.deboi@uantwerpen.be