Development of a Non-Parametric Gaussian Process Model in V3FIT

E. C. Howell, J. D. Hanson

Auburn University, USA

IAEA Technical Meeting on Fusion Data Processing, Validation, and Analysis
Boston, USA
30 May - 2 June, 2017
Outline

- Introduce V3FIT equilibrium reconstruction code
- Review Gaussian process regression
- Discuss development in V3FIT
- Testing using synthetic and experimental data
V3FIT is a tool for three-dimensional equilibrium reconstruction

- Reconstruction determines MHD equilibrium from measurements
- V3FIT uses a parametric representation
  - Equilibrium specified by 2 profiles
  - Additional profiles relate MHD quantities to measurements

V3FIT is used on multiple experiments

CTH

MST
V3FIT is a tool for three-dimensional equilibrium reconstruction

- Reconstruction determines MHD equilibrium from measurements
- V3FIT uses a parametric representation
  - Equilibrium specified by 2 profiles
  - Additional profiles relate MHD quantities to measurements
- Optimal set of parameters minimize $\chi^2$ error

$$\chi^2 = \sum_i \left( \frac{S_i^O - S_i^M(p)}{\sigma_i} \right)^2$$

- Observed signals: $S_i^O$
- Modeled signals: $S_i^M$
- Parameters: $p$

Here we develop a non-parametric model which is used with V3FIT’s standard parametric representation

V3FIT is used on multiple experiments

CTH

MST
Gaussian process regression (GPR) treats unknown profiles as normal random functions (Gaussian processes).
- A Gaussian process is defined by a mean $\mu(x)$ and a covariance $K(x, x')$.
- GPR calculates a posterior distribution using Bayesian analysis.
- Evaluate the posterior at a finite number of points: $f_{*i} = f(x_{*i})$. 

Naturally reproduces complex features with sufficient data.

Nearby points are strongly correlated, preventing overfitting.

Quality of fit improves with additional data.

Parametric Regression
- Quality of fit is limited by choice of a model.
- Models must be designed to capture complex features.
- Overly complex models lead to overfitting.
A non-parametric Gaussian Process model is implemented in V3FIT.

Gaussian process regression (GPR) treats unknown profiles as normal random functions (Gaussian processes).
- A Gaussian process is defined by a mean $\mu(x)$ and a covariance $K(x, x')$.
- GPR calculates a posterior distribution using Bayesian analysis.
- Evaluate the posterior at a finite number of points: $f_{*i} = f(x_{*i})$

**Gaussian Process Regression**
- Quality of fit improves with additional data
- Naturally reproduces complex features with sufficient data
- Nearby points are strongly correlated, preventing overfitting

**Parametric Regression**
- Quality of fit is limited by choice of a model
- Models must be designed to capture complex features
- Overly complex models lead to overfitting
Gaussian process regression is formulated using a Bayesian approach. 

Likelihood for noisy data: \( Y_i = L_i f(x) \)

\[
p(\mathbf{Y}|f) = \prod_i \mathcal{N}(L_i f(x), \sigma_{ni})
\]

Gaussian Process Prior

\[
p(f) = \mathcal{GP}(\mu(x), k(x, x'))
\]

Bayes' Theorem

\[
p(f|\mathbf{Y}) = \frac{p(\mathbf{Y}|f)p(f)}{p(\mathbf{Y})}
\]

Evaluate posterior at specified locations: \( x^* \)

\[
p(f^*|\mathbf{Y}) = p(f(x^*)|\mathbf{Y})
\]
We consider diagnostics that are linearly related to a physical quantity.

- Posterior distribution is a multivariate normal distribution
  \[ p(f_* | Y) \propto \mathcal{N}(m_{f*}, \Sigma_{f*}) \]

  \[
  m_{f*} = L' K (x_*, x) \left( L L' K (x, x') + \Sigma_y \right)^{-1} y \\
  \Sigma_{f*} = K (x_*, x_*) - L' K (x_*, x) \left( L L' K (x, x') + \Sigma_y \right)^{-1} L K (x, x_*)
  \]

- We consider the squared exponential Kernel function with zero mean.
  \[
  K(x, x') = \sigma_f^2 \exp\left(-\frac{(x - x')^2}{2\sigma_L^2}\right), \quad \mu(x) = 0
  \]

---

\(^1\) J. Svensson, EFDA-JET-PR(11)24, 2011
We consider diagnostics that are linearly related to a physical quantity.

Posterior distribution is a multivariate normal distribution\(^1\)

\[
p(f_* \mid Y) \propto \mathcal{N}\left(\mathbf{m}_{f_*}, \Sigma_{f_*}\right)
\]

\[
\mathbf{m}_{f_*} = \mathbf{L}' \mathbf{K} (\mathbf{x}_*, \mathbf{x}) (\mathbf{L} \mathbf{L}' \mathbf{K} (\mathbf{x}, \mathbf{x}') + \Sigma_y)^{-1} \mathbf{y}
\]

\[
\Sigma_{f_*} = \mathbf{K} (\mathbf{x}_*, \mathbf{x}_*) - \mathbf{L}' \mathbf{K} (\mathbf{x}_*, \mathbf{x}) (\mathbf{L} \mathbf{L}' \mathbf{K} (\mathbf{x}, \mathbf{x}') + \Sigma_y)^{-1} \mathbf{L} \mathbf{K} (\mathbf{x}, \mathbf{x}_*)
\]

We consider the squared exponential Kernel function with zero mean.

\[
\mathbf{K} (\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{(\mathbf{x} - \mathbf{x}')^2}{2\sigma_L^2}\right), \quad \mu (\mathbf{x}) = 0
\]

Optimal set of hyperparameters is found by maximizing the evidence

\[
\text{Evidence} = \frac{1}{2} \left( N \ln (2\pi) + \ln |\mathbf{L} \mathbf{L}' \mathbf{K} + \Sigma_y| + \mathbf{y}^T (\mathbf{L} \mathbf{L}' \mathbf{K} + \Sigma_y)^{-1} \mathbf{y} \right)
\]

\(^1\) J. Svensson, EFDA-JET-PR(11)24, 2011

---

E. C. Howell, J. D. Hanson

Development of V3FIT GP Model
The Gaussian process implementation in V3FIT is designed to work in conjunction with standard parametric regression.

- Gaussian processes are used to model profiles that are linearly related to measured data
  - Soft X-ray emissivity profiles
  - $T_e$ from ECE and Thomson
  - $n_e$ from interferometry
- Parametric representation is used for remaining profiles
- $\chi^2$ minimization is used to optimize the parameters
  - A quasi-Newton method is used to find the optimal set of parameters
  - The Gaussian process profiles are recalculated at each iteration
  - Hyperparameters are optimized by maximizing the evidence
The implementation of the Gaussian process is tested using synthetic data. The Gaussian process infers the temperature profile from synthetic Thomson signals. Synthetic Thomson signals are calculated for a prescribed temperature profile. Tests use both pure and noisy signals. The equilibrium flux surfaces are fixed in these tests.
The Gaussian process reproduces the model temperature profile

- The best fit is the mean of the Gaussian Process with optimal hyperparameters
- Samples of the posterior distribution characterize the uncertainty
- Most uncertainty in the edge ($S \gtrsim 0.6$) where measurements are sparse
- This example shown uses synthetic data with 10% noise

E. C. Howell, J. D. Hanson
Development of V3FIT GP Model
A sawtoothing CTH discharge is reconstructed using Gaussian processes to infer the soft X-ray emissivity profiles.

- The Compact Toroidal Hybrid (CTH) is a 5 period current carrying stellarator.
- Reconstruction uses a set $\sim 250$ diagnostic signals:
  - Multiple external magnetic diagnostics
  - 3 two-color soft X-ray cameras
  - 1 one-color soft X-ray camera
  - 3 interferometry chords
- Measured soft X-ray signals constrain the shape of internal flux surfaces:
  - Soft X-ray emissivity profiles are assumed to be a flux-functions
  - Each soft X-ray color is has a unique emissivity profile
  - Each emissivity profiles is inferred using a Gaussian Process
Hybrid reconstructions that use Gaussian processes agree with fully parametric reconstruction

**Hybrid Reconstruction**
- 3 Gaussian process emissivity profiles
- Parametric density, current, and pressure profiles
- 13 reconstructed parameters and hyperparameters

**Parametric Reconstruction**
- 3 parametric emissivity profiles
- Parametric density, current, and pressure profiles
- 37 reconstructed parameters
A MST helical axis discharge is used as a second test case

- Madison Symmetric Torus (MST) bifurcates to a helical equilibrium at high current\(^2\)
- Reconstructions represent the temperature profile using a Gaussian process
- Reconstructions use approximately 200 diagnostics
  - Multiple external magnetic diagnostics
  - 4 soft X-ray cameras
  - Thomson scattering
  - FIR interferometry/polarimetry

---
- Flux surfaces overlap (previous slide)
- Reconstructed temperature profiles agree within error bars
  - Parametric temperature profile shows signs of overfitting
  - Overfitting can be addressed by reducing the number of knots in the spline
  - The Gaussian process automatically addresses overfitting
A hybrid regression model has been implemented in V3FIT
- Gaussian process regression is used to infer select profiles
- Standard parametric regression is used to infer the remaining profiles and other free parameters

The Gaussian process model accurately reproduces the profile when synthetic data is used

Tests using experimental data shows good agreement between hybrid reconstructions and fully parametric reconstructions

Modular programming makes it easy to incorporate new kernels and additional functionality
References