Schedule

- 16 Jan: Gaussian processes (Jo Eidsvik)
- **23 Jan: Hands-on project on Gaussian processes (Team effort, work in groups)**
- 30 Jan: Latent Gaussian models and INLA (Jo Eidsvik)
- 6 Feb: Hands-on project on INLA (Team effort, work in groups)
- 12-13 Feb: Template model builder. (Guest lecturer Hans J Skaug)
- To be decided...

Project A: Bayesian optimization

Gaussian processes are commonly used in optimization of complex functions.

Usually the function $Y(t)$ is very expensive to evaluate.

Goal

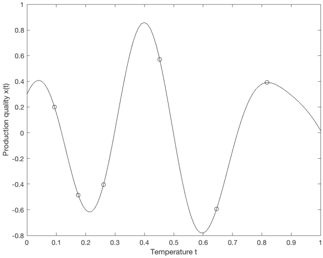
$$\hat{t} = \text{argmax} Y(t)$$

Example: t is temperature, $Y(t)$ is production quality. Which temperature input gives highest quality?

Project A: Bayesian optimization

In this project the function is simply

$$Y(t) = 0.5\cos(15t) + 0.25\sin(20t) - 0.2\cos(10t)$$



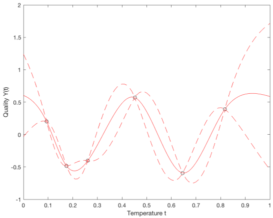
Project A: Bayesian optimization

A set is a grid from 0.00 to 1.00, with step size 0.01.
B set $(0.094, 0.175, 0.261, 0.4520, 0.6460, 0.817)$, with measurements $(0.2002, -0.4868, -0.4038, 0.5716, -0.5928, 0.3904)$.

$$\begin{aligned} E(\boldsymbol{Y}_A|\boldsymbol{Y}_B) &= \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{A,B}\boldsymbol{\Sigma}_B^{-1}(\boldsymbol{Y}_B - \boldsymbol{\mu}_B), \\ \mathrm{Var}(\boldsymbol{Y}_A|\boldsymbol{Y}_B) &= \boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{A,B}\boldsymbol{\Sigma}_B^{-1}\boldsymbol{\Sigma}_{B,A}. \end{aligned}$$

Model

$$\mu(t) = -0.05, \quad \Sigma(s,t) = (1 + 6|t-s|)\exp(-6|t-s|).$$



Project A: Bayesian optimization

Expected improvement:

$$\begin{aligned} \mathrm{EI} &= E(\max\{0, Y(t) - Y^*\} | \boldsymbol{Y}_B) \\ &= (\hat{\mu}(t) - Y^*)\Phi\left[\frac{\hat{\mu}(t) - Y^*}{\hat{\sigma}(t)}\right] + \hat{\sigma}(t)\phi\left[\frac{\hat{\mu}(t) - Y^*}{\hat{\sigma}(t)}\right] \end{aligned}$$

$$Y^* = \max \boldsymbol{Y}_B$$

$\hat{\mu}(t)$ and $\hat{\sigma}(t)$ are posterior mean and standard deviation, given $\boldsymbol{Y}_B$.
Φ and ϕ is cdf and pdf of standard Gaussian distribution.

Project A: Bayesian optimization

Sequential optimization using expected improvement.
Repeat the following for some iterations:

- Use EI to find next best point, given current data.
- Evaluate next point.
- Augment B set with this observation.

Project B: Spatial regression model

Model: $Y(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\beta + w(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$.

1. $Y(\boldsymbol{s})$ response variable at location $\boldsymbol{s} = (s_1, s_2)$.
2. β regression effects. $\boldsymbol{X}(\boldsymbol{s})$ covariates at $\boldsymbol{s}$.
3. $w(\boldsymbol{s})$ structured (spatially correlated) Gaussian process.
4. $\epsilon(\boldsymbol{s})$ unstructured (independent) Gaussian measurement noise.

Project B: Spatial regression model

True model:

- $\boldsymbol{s}$ is in the unit square.
- $Y(\boldsymbol{s}) = \beta + w(\boldsymbol{s}) + N(0, \tau^2)$, $\beta = 1$, $\tau^2 = 0.1^2$.
- Exponential covariance: $\text{Cov}(w(\boldsymbol{s}), w(\boldsymbol{t})) = \sigma^2 \exp(-\phi|\boldsymbol{s} - \boldsymbol{t}|)$, $\sigma^2 = 1^2$, $\phi = 6$.

Goal is to estimate parameters β and $\boldsymbol{\theta} = (\sigma^2, \phi, \tau^2)$ from data at $n = 100$ random locations: $\boldsymbol{Y} = (Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_n))'$.

Project B: Spatial regression model

MLE:

$$(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\beta}}) = \text{argmax}_{\boldsymbol{\theta},\boldsymbol{\beta}}\{l(\boldsymbol{Y}; \boldsymbol{\beta}, \boldsymbol{\theta})\}.$$

Distance matrix $\boldsymbol{H}$ of size 100×100.

$$\boldsymbol{Z} = (\boldsymbol{Y} - \underline{1}\beta), \boldsymbol{\Sigma} = \sigma^2 \exp(-\phi \boldsymbol{H}), \boldsymbol{C} = \boldsymbol{\Sigma} + \tau^2 \boldsymbol{I}, \boldsymbol{Q} = \boldsymbol{C}^{-1}.$$

$$\frac{dl}{d\boldsymbol{\theta}_r} = -\frac{1}{2}\text{trace}(\boldsymbol{Q}\frac{d\boldsymbol{C}}{d\theta_r}) + \frac{1}{2}\boldsymbol{Z}'\boldsymbol{Q}\frac{d\boldsymbol{C}}{d\theta_r}\boldsymbol{Q}\boldsymbol{Z},$$

$$E\left(\frac{d^2 l}{d\theta_r d\theta_s}\right) = -\frac{1}{2}\text{trace}(\boldsymbol{Q}\frac{d\boldsymbol{C}}{d\theta_s}\boldsymbol{Q}\frac{d\boldsymbol{C}}{d\theta_r}).$$

$$\frac{d\boldsymbol{C}}{d\sigma^2} = \exp(-\phi \boldsymbol{H}), \qquad \frac{d\boldsymbol{C}}{d\phi} = -\sigma^2 \boldsymbol{H} \odot \exp(-\phi \boldsymbol{H}), \qquad \frac{d\boldsymbol{C}}{d\tau^2} = \boldsymbol{I}.$$

Project B: Spatial regression model

$$\boldsymbol{Q} = \boldsymbol{Q}(\theta_p)$$

$$\hat{\beta}_p = [\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X}]^{-1}\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{Y},$$

$$\hat{\boldsymbol{\theta}}_{p+1} = \hat{\boldsymbol{\theta}}_p - E\left(\frac{d^2 l(\boldsymbol{Y}; \hat{\beta}_p, \hat{\boldsymbol{\theta}}_p)}{d\boldsymbol{\theta}^2}\right)^{-1} \frac{dl(\boldsymbol{Y}; \hat{\beta}_p, \hat{\boldsymbol{\theta}}_p)}{d\boldsymbol{\theta}},$$

Project B: Spatial regression model

Iterative scheme can here start with true β, σ^2, ϕ and τ^2.
Sample $n = 100$ random sites in the unit square. Form the distance matrix, the covariance matrix, and its Cholesky factor based on the true model.
Repeat the following for some Monte Carlo realizations:

- Draw a realization $\boldsymbol{Y}$.
- Find MLE of the model parameters. Monitor convergence over iterations $p = 1, \ldots, 10$.

The estimates should be close to the true parameters, with some variability over the realizations.