Extending Model-based Policy Gradients for Robots in Hetroscedastic Environments

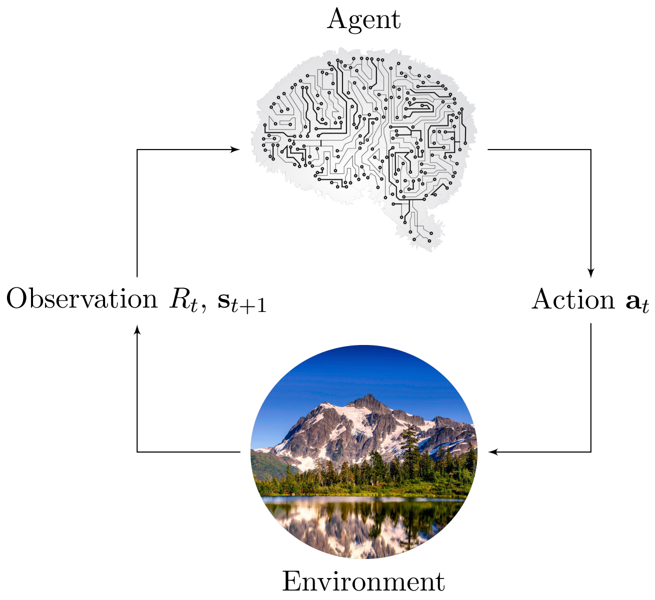
John Martin Jr. and Brendan Englot
Stevens Institute of Technology

October 6, 2017

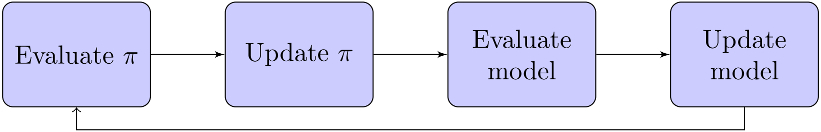
Context

- Enable robust autonomy in marine robots
- The marine environment exhibits complex behavior
- Sensor performance depends on the environment
- Learn complex behavior from repeated operation

Reinforcement Learning



Model-based Reinforcement Learning



- Evaluate π with model rollouts
- Update policy parameters in gradient direction
- Evaluate model with experience
- Update model for better predictions

Gaussian Process Regression

- Statistical model is $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}$
- $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_{ff})$, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$

Gaussian Process Regression

- Statistical model is $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}$
- $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_{ff})$, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$
- Gaussian distributions fully specified by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$
- Gaussian processes fully specified by $\mu(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$

Gaussian Process Regression

- Statistical model is $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}$
- $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}_{ff})$, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$
- Gaussian distributions fully specified by $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$
- Gaussian processes fully specified by $\mu(\mathbf{x})$ and $k(\mathbf{x}, \mathbf{x}')$

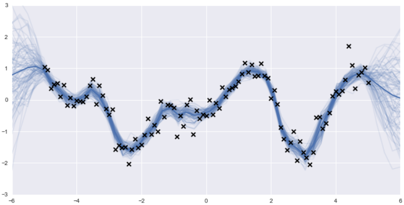
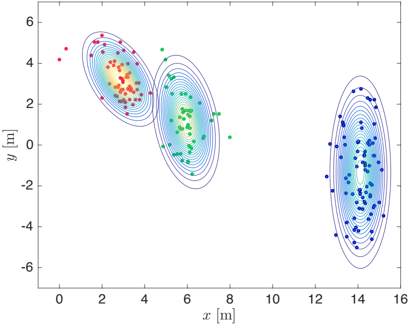


Figure: Image from this tutorial.

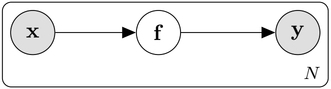
Heteroscedastic Observations



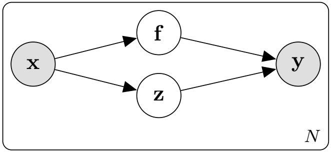
- Noise levels have a functional dependence on $\mathbf{x}$
- Constant noise levels no longer valid

$$\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}) \qquad \boldsymbol{\varepsilon}(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, z(\mathbf{x})\mathbf{I})$$

Heteroscedastic Gaussian Processes



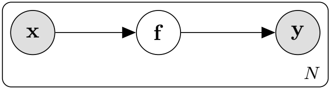
(a) Standard



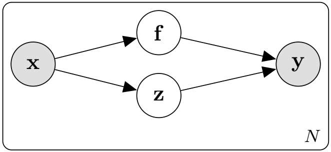
(b) Heteroscedastic

- Statistical model becomes $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}(\mathbf{x})$

Heteroscedastic Gaussian Processes



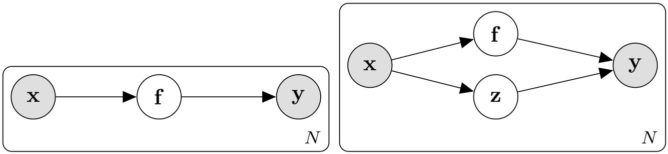
(a) Standard



(b) Heteroscedastic

- Statistical model becomes $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}(\mathbf{x})$
- What process specifies $\boldsymbol{\varepsilon}(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, z(\mathbf{x})\mathbf{I})$?

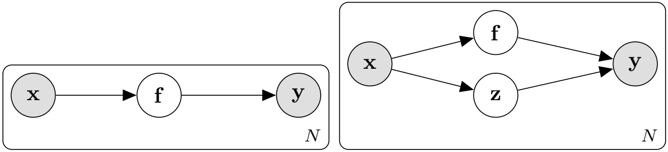
Heteroscedastic Gaussian Processes



(a) Standard (b) Heteroscedastic

- Statistical model becomes $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}(\mathbf{x})$
- What process specifies $\boldsymbol{\varepsilon}(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, z(\mathbf{x})\mathbf{I})$?
- $z(\mathbf{x}) \sim \exp[\mathcal{N}(0, \sigma^2\mathbf{I})]$ [Goldberg et al., 1997]

Heteroscedastic Gaussian Processes



(a) Standard (b) Heteroscedastic

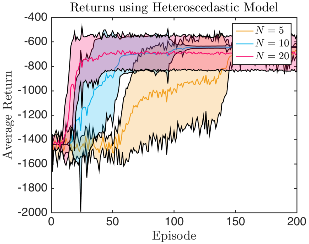
- Statistical model becomes $\mathbf{y} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varepsilon}(\mathbf{x})$
- What process specifies $\boldsymbol{\varepsilon}(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, z(\mathbf{x})\mathbf{I})$?
- $z(\mathbf{x}) \sim \exp[\mathcal{N}(0, \sigma^2\mathbf{I})]$ [Goldberg et al., 1997]

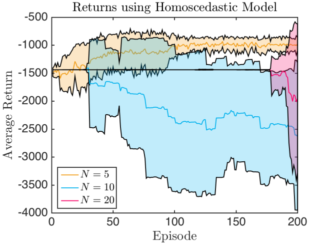
$$p(y_*|\mathbf{x}_*, \mathbf{X}, \mathbf{y}) = \int p(\mathbf{y}|x_*, \mathbf{X}, \mathbf{y}, \mathbf{z}, z_*)p(\mathbf{z}, z_*|\mathbf{X}, \mathbf{y})d\mathbf{z}dz_*$$

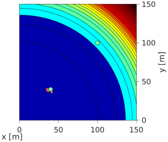
- Predict with EM-style algorithm [Kersting et al., 2007]

$$p(y_*|\mathbf{x}_*, \mathbf{X}, \mathbf{y}) \approx p(\mathbf{y}|x_*, \mathbf{X}, \mathbf{y}, \mathbf{z}^*, z_*^*)$$

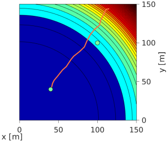
Results



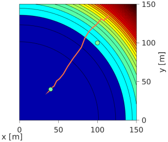




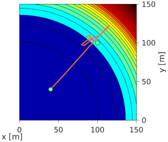
(a) 50 Ep



(b) 100 Ep



(c) 150 Ep



(d) 200 Ep

Questions

References I

[0] Goldberg, P., Williams, C., and Bishop, C. (1997).
Regression with input-dependent noise: A gaussian process treatment.
In *Proceedings of the 10th International Conference on Neural Information Processing Systems*.

[0] Kersting, K., Plagemann, C., Pfaff, P., and Burgard, W. (2007).
Most likely heteroscedastic gaussian process regression.
In *Proceedings of the 24th International Conference on Machine Learning*.