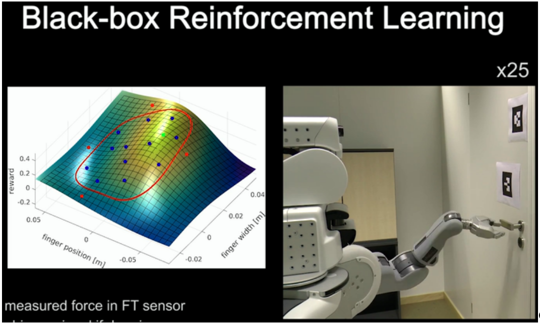
Applications of Constrained BayesOpt in Robotics and Rethinking Priors & Hyperparameters

Marc Toussaint

Machine Learning & Robotics Lab – University of Stuttgart
marc.toussaint@informatik.uni-stuttgart.de

NIPS BayesOpt, Dec 2016

(1) Learning Manipulation Skills



Englert & Toussaint: *Combined Optimization and Reinforcement Learning for Manipulation Skills*. R:SS'16

Combined Black-Box and Analytical Optimization

Englert & Toussaint: *Combined Optimization and Reinforcement Learning for Manipulation Skills*. R:SS'16

- CORL (Combined Optimization and RL):
  - Policy parameters w
  - **analytically known cost function** $J(w) = \mathsf{E}\{\sum_{t=0}^T c_t(x_t, u_t) \,|\, w\}$
  - **projection**, implicitly given by a constraint $h(w, \theta) = 0$
  - **unknown black-box return function** $R(\theta) \in \mathbb{R}$
  - **unknown black-box success constraint** $S(\theta) \in \{0, 1\}$
  - Problem:

$$\min_{w,\theta} J(w) - R(\theta) \quad \text{s.t.} \quad h(w, \theta) = 0,\ S(\theta) = 1$$

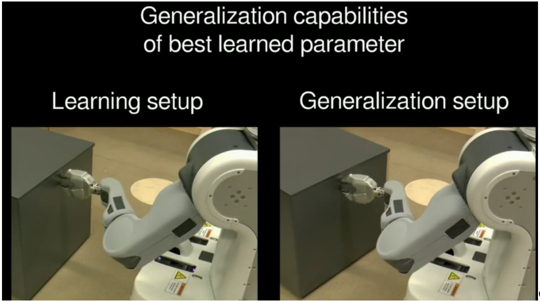
- Alternate path optimization $\min_w J(w)$ s.t. $h(w, \theta) = 0$
  with Bayesian Optimization $\max_\theta R(\theta)$ s.t. $S(\theta) = 1$

Heuristic to handle constraints

- Prior mean $\mu = 2$ for g
- Sample only points s.t. $g(x) \leq 0$
- Acquisition function combines PI with Boundary Uncertainty

$$\alpha_{\mathsf{PIBU}}(x) = [g(x) \geq 0]\mathsf{PI}_f(x) + [g(x) = 0]\beta\sigma_g^2(x)$$

(2) Optimizing Controller Parameters



Drieß, Englert & Toussaint: *Constrained Bayesian Optimization of Combined Interaction Force/Task Space Controllers for Manipulations*. IROS Workshop'16

Controller Details

- Non-switching controller for smoothly establishing contacts
  - In (each) task space

$$\ddot{y}^* = \ddot{y}^{\text{ref}} + K_p(y^{\text{ref}} - y) + K_d(\dot{y}^{\text{ref}} - \dot{y})$$

  - Operational space controller (linearized)

$$\begin{aligned}
\ddot{q}^* &= \bar{K}_p q + \bar{K}_d \dot{q} + \bar{k} \\
\bar{K}_p &= (H + J^\top C J)^{\text{-1}} [H K_p^q + J^\top C K_p J] \\
\bar{K}_d &= (H + J^\top C J)^{\text{-1}} [H K_d^q + J^\top C K_d J] \\
\bar{k} &= (H + J^\top C J)^{\text{-1}} [H k^q + J^\top C k]
\end{aligned}$$

  - Contact force limit control

$$e \leftarrow \gamma e + [|f| > |f^{\text{ref}}|] \, (f^{\text{ref}} - f) \qquad\qquad u = J^\top \alpha e$$

- **Many parameters!** Esp. $\alpha, \dot{y}^{\text{ref}}, K_d$

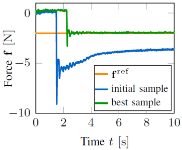
Optimizing Controller Parameters

- Optimization objectives:
  - Low compliance: $\mathrm{tr}(\bar{K}_p)$ and $\mathrm{tr}(\bar{K}_d)$
  - Contact force error: $\int (f^{\text{ref}} - f)^2 \, dt$
  - Peak force on onset: $|f^{os}|$
  - Smooth force profile: $\int \left(|\frac{d}{dt} f(t)| + |\frac{d^2}{dt^2} f(t)| + |\frac{d^3}{dt^3} f(t)| \right) dt$
  - Boolean success: contact and *staying in* contact

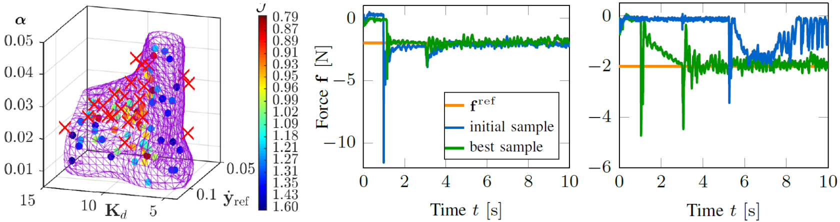
Optimizing Controller Parameters

- Optimization objectives:
  - Low compliance: $\mathrm{tr}(\bar{K}_p)$ and $\mathrm{tr}(\bar{K}_d)$
  - Contact force error: $\int (f^{\mathrm{ref}} - f)^2 \; dt$
  - Peak force on onset: $|f^{os}|$
  - Smooth force profile: $\int \left(|\frac{d}{dt} f(t)| + |\frac{d^2}{dt^2} f(t)| + |\frac{d^3}{dt^3} f(t)| \right) \; dt$
  - Boolean success: contact and *staying in* contact

- Establishing contact

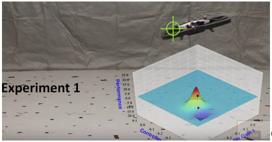


- Sliding



(3) Safe Active Learning & BayesOpt

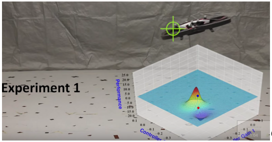
- SafeOpt: Safety threshold on the objective $f(x) \geq h$



Sui, Gotovos, Burdick, Krause: *Safe Exploration for Optimization with Gaussian Processes*. ICML'15

(3) Safe Active Learning & BayesOpt

- SAFEOPT: Safety threshold on the objective $f(x) \geq h$

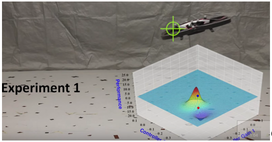


Sui, Gotovos, Burdick, Krause: *Safe Exploration for Optimization with Gaussian Processes*. ICML'15

- Guarantee to never step outside an unknown $g(x) \leq 0$...
  - Impossible when no failure data $g(x) > 0$ exists...

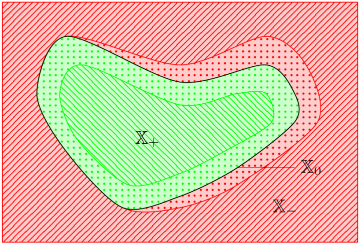
(3) Safe Active Learning & BayesOpt

- SAFEOPT: Safety threshold on the objective $f(x) \geq h$



Sui, Gotovos, Burdick, Krause: *Safe Exploration for Optimization with Gaussian Processes*. ICML'15

- Guarantee to never step outside an unknown $g(x) \leq 0$...
  - Impossible when no failure data $g(x) > 0$ exists...
  - Unless you assume observation of near boundary discriminative values



Schreiter et al: *Safe Exploration for Active Learning with Gaussian Processes*. ECML'15

Probabilistic guarantees on non-failure

- Acquisition function

$$\alpha(x) = \sigma_f^2(x) \quad \text{s.t.} \quad \mu_g(x) + \nu\sigma_g(x) \geq 0$$

- Specify probability of failure δ after n points with m_0 initializations $\mapsto \nu$
- Application on cart-pole

δ	ν	m_0	H	# failures	# expected failures
0.01	4.26	8	393.3	0	9.9
0.05	3.89	7	397.7	6	48.4
0.10	3.72	6	412.8	29	94.6
0.20	3.54	6	402.7	57	180.2
0.30	3.43	5	395.6	80	257.9
0.40	3.35	5	389.3	101	328.1
0.50	3.29	5	380.7	127	391.6

So, what are the issues?

So, what are the issues?

– Choice of hyper parameters!

So, what are the issues?

- Choice of hyper parameters!
- Stationary covariance functions!

So, what are the issues?

- Choice of hyper parameters!
- Stationary covariance functions!
- Isotropic stationary covariance functions!

- Actually, I’m a fan of Newton Methods

- Actually, I'm a fan of Newton Methods
- Two messages of classical (convex) optimization:
  - Step size (line search, trust region, Wolfe)
  - Step direction (Newton, quasi-Newton, BFGS, conjugate, covariant)
- Newton methods are prefect for local optimum down-hill

Model-based Optimization

- If the model is not given: classical *model-based optimization* (Nodecal et al. “Derivative-free optimization”)

Initialize D with at least $\frac{1}{2}(n+1)(n+2)$ data points
repeat
    Compute a regression $\hat{f}(x) = \phi_2(x)^\top \beta$ on D
    Compute $x^+ = \operatorname{argmin}_x \hat{f}(x)$ s.t. $|x - \hat{x}| < \alpha$
    Compute the improvement ratio $\varrho = \frac{f(\hat{x}) - f(x^+)}{\hat{f}(\hat{x}) - \hat{f}(x^+)}$
    if $\varrho > \epsilon$ **then**
        Increase the stepsize α
        Accept $\hat{x} \leftarrow x^+$
        Add to data, $D \leftarrow D \cup \{(x^+, f(x^+))\}$
    else
        if $\det(D)$ is too small **then** *// Data improvement*
            Compute $x^+ = \operatorname{argmax}_x \det(D \cup \{x\})$ s.t. $|x - \hat{x}| < \alpha$
            Add to data, $D \leftarrow D \cup \{(x^+, f(x^+))\}$
        else
            Decrease the stepsize α
        end if
    end if
    Prune the data, e.g., remove $\operatorname{argmax}_{x \in \Delta} \det(D \setminus \{x\})$
until x converges

This is similar to BayesOpt with polynomial kernel!

A prior about local polynomial optima

- Assume that the objective has multiple local optima
  - Local optimum: locally convex
  - Each local optimum might be differently conditioned

  $\rightarrow$ we need a highly non-stationary, non-isotropic converance function
- "Between" the local optima, the function is smooth $\rightarrow$ standard squared exponential kernel

A prior about local polynomial optima

- Assume that the objective has multiple local optima
  - Local optimum: locally convex
  - Each local optimum might be differently conditioned

  $\rightarrow$ we need a highly non-stationary, non-isotropic converance function
- "Between" the local optima, the function is smooth $\rightarrow$ standard squared exponential kernel
- The **Mixed-global-local kernel**

$$k_{\text{MGL}}(x, x') = \begin{cases} k_q(x, x'),\ x, x' \in \mathcal{U}_i, \\ k_s(x, x'), x \notin \mathcal{U}_i, x' \notin \mathcal{U}_j \\ 0,\ \text{else} \end{cases}$$

for any i, j

$$k_q(x, x') = (x^T x' + 1)^2$$

Finding convex neighborhoods

- Data set $D = \{(x_i, y_i)\}$
- $\mathcal{U} \subset \mathcal{D}$ is a convex neighborhood if

$$\{\beta_0^*, \beta^*, B^*\} = \underset{\beta_0, \beta, B}{\operatorname{argmin}} \sum_{k : x_k \in \mathcal{U}} \left[(\beta_0 + \beta^T x_k + \frac{1}{2} x_k^T B x_k) - y_k \right]^2$$

has a positive definite Hessian B^*

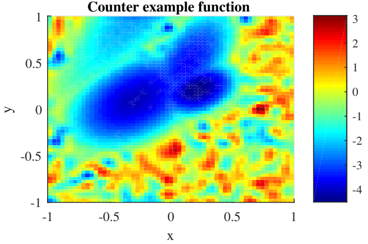
A heuristic to decrease length-scale

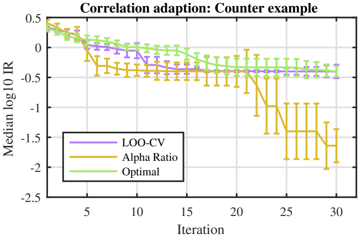
- The SE-part still has a length-scale hyperparameter l
- In each iteration we *consider* to decrease to $\tilde{l}_t < l_{t-1}$

$$\alpha_{r,t} := \frac{\alpha^*(\tilde{l}_t)}{\alpha^*(l_{t-1})}\ , \quad \alpha^*(l) = \min_x \alpha(x; l)$$

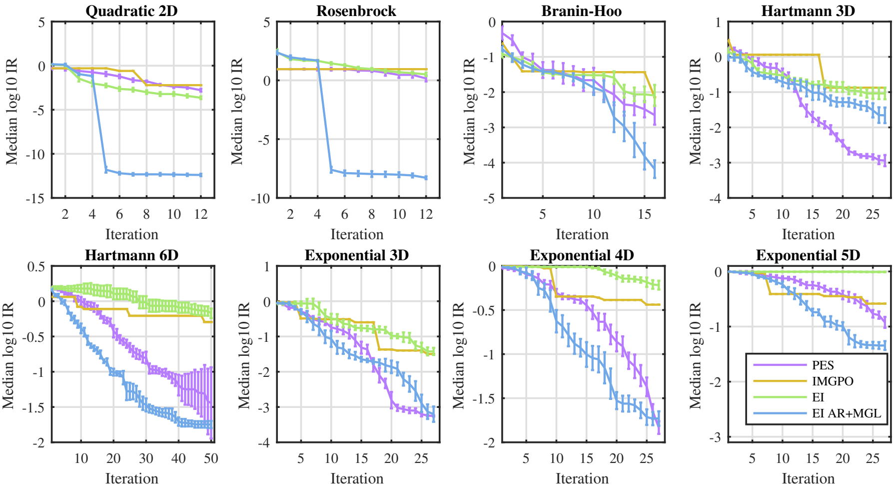
  for any acquisition function $\alpha(x; l)$

- Accept smaller lengthscale only if $\alpha_{r,t} \geq h$ (e.g., $h \approx 2$)
- Robust to non-stationary objectives





Mixed-global-local kernel + alpha ratio



- PES: Bayesian integration over hyper parameters
- IMGPO: Bayesian update for hyperparameters in each iteration

...work with **Kim Wabersich**

Conclusions

- Solid optimization methods are the savior of robotics!
- Rethink the priors we use for BayesOpt
  - Local optima with varying conditioning
- Rethink the objective for choosing hyper parameters
  - Maximize optimization progress ($\sim$ expected acquisition) rather than data likelihood

Thanks

- *for your attention!*
- to the students:
  - Peter Englert (BayesOpt for Manipulation)
  - Jens Schreiter (Safe Active Learning)
  - Danny Drieß(BayesOpt for Controller Optimization)
  - Kim Wabersich (Mixed-global-local kernel & alpha ratio)
- and my lab: