ETHz PAI Cheat Sheet

Probability & BLR

- Gaussian

$P(\mathcal{N}|\mathcal{N}) \sim \mathcal{N}(\mu_{A|B}, \Sigma_{A|B})$

$$\mu_{A|B} = \mu_A + \Sigma_{AB}\Sigma_{BB}^{-1}(x_B - \mu_B)$$
$$\Sigma_{A|B} = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$$

$P(M\mathcal{N}) \sim \mathcal{N}(\mu_Y, \Sigma_Y)$

$$\mu_Y = M\mu_X$$
$$\Sigma_{YY} = M^T \Sigma_{XX} M$$

$P(\mathcal{N} + \mathcal{N}) \sim P(\mu_Y, \Sigma_Y)$

$$\Sigma_{YY} = (\Sigma_{XX}^{-1} + \Sigma_{X'X'}^{-1})^{-1}$$
$$\mu_Y = \Sigma_{YY}\Sigma_{XX}^{-1}\mu_X + \Sigma_{YY}\Sigma_{X'X'}^{-1}\mu_{X'}$$

$P(\mathcal{N}\mathcal{N}) \sim P(\mu_Y, \Sigma_Y)$

$$\Sigma_{YY} = (\Sigma_{XX}^{-1} + \Sigma_{X'X'}^{-1})^{-1}$$
$$\mu_Y = \Sigma_{YY}\Sigma_{XX}^{-1}\mu_X + \Sigma_{YY}\Sigma_{X'X'}^{-1}\mu_{X'}$$

entropy: $ln(\sigma\sqrt{2\pi e})$ (max in all distribution)

- KL-Divergence:

$$D_{KL}(p||q) = \int p(x) log \frac{p(x)}{q(x)} dx = H(p|q) - H(p)$$

- $D_{KL}(p||q)$: mode averageing
- $D_{KL}(q||P)$: mode seeking
  $p \sim \mathcal{N}(0, diag(\sigma_1^2, \sigma_2^2)) \Rightarrow \sigma_q^2 = \frac{2}{\sigma_1^{-2} + \sigma_2^{-2}}$
- $D_{KL}(q||p)$ is well defined if q is a subset of p
- $q \sim \mathcal{N}(\mathbb{E}(p, Var(p)) \Rightarrow H(p|q) = H(q)$

- Bayesian Linear Regression:

$$y = w^\top x + \epsilon \quad \epsilon \sim \mathcal{N}(0, \sigma_n^2) \; w \sim \mathcal{N}(0, \sigma_p^2)$$
$$P(w|Y, X) \sim \mathcal{N}(\mu, \Sigma)$$
$$\mu = \frac{1}{\sigma_n^2}\Sigma X^T Y \quad \Sigma = \left(\frac{1}{\sigma_n^2} X^T X + \frac{1}{\sigma_p^2} I\right)^{-1}$$

uncertainty: aleartoric + epistemic

Online

$$X_{new} X^\top X_{new} = X^T X + x_{t+1} x_{t+1}^\top$$
$$X_{new}^\top Y_{new} = X^T Y + y_{t+1} x_{t+1}$$

Fast : $\mathcal{O}(d^3) \to \mathcal{O}(d^2)$

$$(A + xx^\top)^{-1} = A^{-1} - \frac{(A^{-1}x)(A^{-1}x)^T}{1 + x^\top A^{-1} x}$$

$$(X_{new}^\top X_{new} + \sigma_n^2 I)^{-1} = (\underbrace{X^\top X + \sigma_n^2}_{A} + x_{t+1} x_{t+1}^\top)^{-1}$$

GP & Kalman Filter

- Gaussian Process

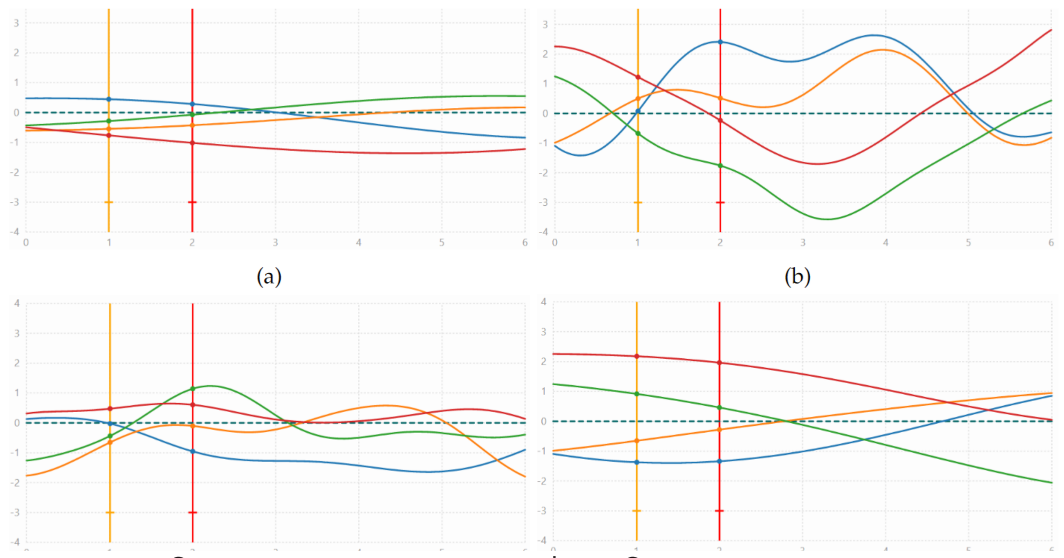
$$y = f(x) + \epsilon \quad \epsilon \sim \mathcal{N}(0, \sigma_n^2)$$
$$P(f|X, Y) \sim GP(f; \mu', k')$$
$$\mu'_{x^*} = \mu_{x^*} + K_{x^* X}(K_{XX} + \sigma_n^2 I)^{-1} Y$$
$$k'_{x^* x^*} = K_{x^* x^*} - K_{x^* X}^T (K_{XX} + \sigma_n^2 I)^{-1} K_{x^* X}$$

Kernel

RBF
$$k(u, v) = \sigma_F^2 exp\left(-\frac{(u - v)^2}{2l^2}\right)$$

l: length scale control the distance of data
σ_F : output scale control the magnitude

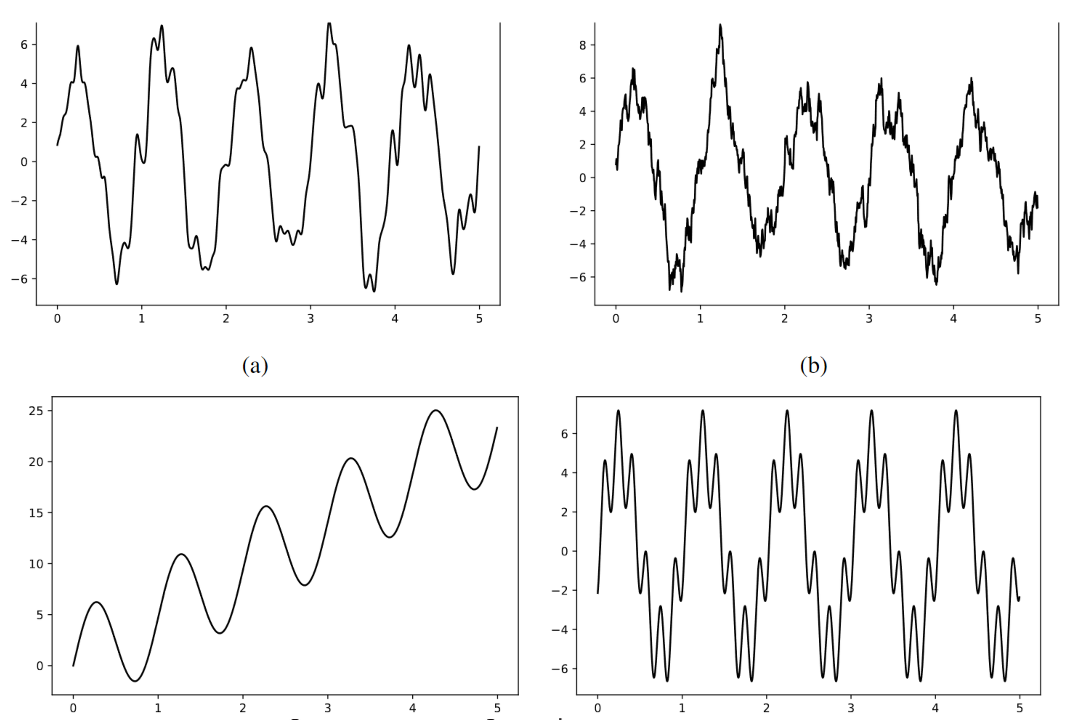
$$\lim_{l \to 0} k(u, v) = \sigma_F^2 \delta(u - v)$$
$$\lim_{l \to 0} k'(x, x) = k(x, x) - k_{xX}(K_{XX} + \sigma_n^2 I)^{-1} k_{xX}^\top$$
$$= \frac{\sigma_F^2 \sigma^2}{\sigma_F^2 + \sigma^2}$$



$\sigma_F^2 = 0.5 \quad l = 4$	$\sigma_F^2 = 2 \quad l = 1$
$\sigma_F^2 = 0.5 \quad l = 1$	$\sigma_F^2 = 2 \quad l = 4$

Linear
$$k(u, v) = \lambda uv$$

equals to BLR with $\lambda = \sigma_p^2$



$exp(-\lambda^2(x - y)^2)$	$exp(-\lambda\|x - y\|)$
λxy	$cos(2\lambda\|x - y\|)$

GP & Kalman Filter

Fast

- $k(x, x') = \phi(x)\phi(x)^\top \; \mathcal{O}(n^3) \to \mathcal{O}(nm^2 + m^3)$
- fourier features: $k(x, x') \approx k(x - x')$
  $= \int_{\mathbb{R}^d} p(\omega) e^{j\omega^T(x - x')} dw$
  Bochner Theorem:$p(\omega) \geq 0 \Rightarrow k \geq 0$
- inducing points: $\mathcal{O}(n^3)$
  SoR: subset of regressor
  FITC: Fully independent training conditional

- Kalman Filter

$$x_{t+1} = F_t x_t + \Sigma_{x,t} \quad y_t = H_t x_t + \Sigma_y$$

$P(x_{t+1}|x_t) \sim \mathcal{N}(Fx_t, \Sigma_{x,t}) \; P(y_t|x_t) \sim \mathcal{N}(H_t x_t, \Sigma_y)$

predict

$$\hat{x}_{t+1} = F_t x_t \quad \hat{\Sigma}_{x,t+1} = H_t \Sigma_{x,t} H_t^\top + \Sigma_{d_t}$$
$$P(x_t|y_{1:t}) \sim \mathcal{N}(\mu_t, \Sigma_{d_t}^2)$$

correct

$$K_{t+1} = \hat{\Sigma}_{x,t} H_t^\top (H_t \hat{\Sigma}_{x,t} H_t^\top + \Sigma_y)^{-1}$$
$$x_{t+1} = \hat{x}_{t+1} + K_{t+1}(y_{t+1} - H_t \hat{x}_{t+1})$$
$$\Sigma_{x,t+1} = (I - K_{t+1} H_t)\hat{\Sigma}_{x,t}$$

kalman gain:K_{t+1}
assume $P(x_{t+1}|x_t) \sim \mathcal{N}(x_t, \sigma_x^2) \quad x_0 \sim \mathcal{N}(0, \sigma_0^2)$, the Kalman Filter can be seen as a GP
$f \sim \mathcal{GP}(0, k_{KF}) \quad k_{KF}(x, x') = \sigma_0^2 + \sigma_x^2 min\{x, x'\}$

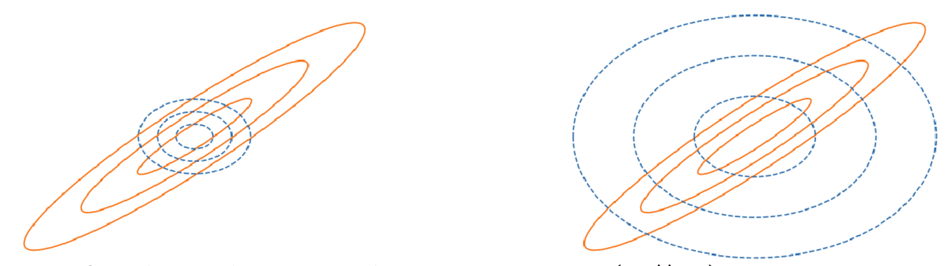
Approximation & MCMC

- Approximation

Laplacian approximation

$q(\theta) \sim (\hat{\theta}, \Lambda^{-1}) \; \hat{\theta} = \underset{\theta}{argmax} p(\theta|y) \; \Lambda = -\nabla\nabla p(\hat{\theta}|y)$

- only one modal
- no previous knowledge



- Left: backward KL $D_{KL}(q||p)$
- Right: forward KL $D_{KL}(p||q)$

Evidence Lower Bound(ELBO)

$$Q^* = \underset{Q \in \mathcal{Q}}{argmin} \, D_{KL}(Q(Z)||P(Z|X))$$
$$= \underset{Q \in \mathcal{Q}}{argmax} \, \mathbb{E}_{Z \sim Q(Z)}[log(P(X|Z))]$$
$$- D_{KL}(Q(Z)||P(Z))$$

MLE $\quad \hat{\theta} = argmax_\theta P(X_{1:n}|\theta)$

sometimes will add l_2 norm which is $\|\cdot\|_2$ for vector and $\|\cdot\|_F$ for matrix

MAP $\quad \hat{\theta} = argmax_\theta (X_{1:n}|\theta, X')$

- Markov Chain Monte Carlo

$$\pi = \pi P$$

where π is the stationary state, P is the transition matrix
ergodic:$\exists t \in \mathbb{N} \to (P)^t > 0$

Metropolis-Hastings (MH) Algorithm:
given proposal transition $R(x'|x)$ and unnormalized stationary distribution $Q(x)$
accept rate $\alpha = min\{1, \frac{Q(x')R(x|x')}{Q(x)R(x'|x)}\}$ from x to x'

Gibbs Sampling : random choose dimension
MALA:
$$R(x'|x) = \mathcal{N}(x'|x - \tau\nabla f(x); 2\tau I)$$

Stochastic Gradient Langevin Dynamics(SGLD)

$$\Delta\theta = -\eta(\nabla log \, p(\theta_t) + \frac{N}{n}\sum_{j=1}^{n} \nabla log p(y_{i_j}|\theta_t, x_{i_j})) + \epsilon_t$$
$$= -\eta(\theta_t + L(\theta_t)) + \epsilon_t$$
$$\epsilon_t \sim \mathcal{N}(0, 2\eta I)$$

where $L(\theta_t)$ is the likelihood
can guarantee convergence.

ETHz PAI Cheat Sheet

Bayesian DL & Optimization

- Bayesian DeepLearning

$$P(y|x,\theta) \sim \mathcal{N}(y|f_\mu(x;\theta_\mu), exp(f_\sigma(x;\theta_\sigma)))$$

train

$$\hat{\theta} = \underset{\theta}{argmin} - lnP(\theta) - \sum_{i=1}^{n} lnP(y_i|x_i,\theta)$$

$$= \underset{\theta}{argmin}\ \lambda||\theta||_2^2 + \frac{1}{2}\sum_{i=1}^{n}\left[\frac{||y_i - \mu(x_i;\theta_\mu)||^2}{\sigma(x_i;\theta_\sigma)^2} + ln\ \sigma(x_i;\theta_\sigma)^2\right]$$

predict

$$P(y'|x',X,Y) = \mathbb{E}_{\theta\sim q}[P(y'|x',\theta)]$$

$$Var[y'|x',X,Y] = \mathbb{E}_{\theta\sim q}[Var[y'|x',\theta]] + Var[\mathbb{E}_{\theta\sim q}[y'|x',\theta]]$$

- Aleatoric uncertainty(random)
- Epistemic uncertainty(knowledge)

- MAP for BNN is not closed
- MLE for BNN is closed

Stochastic Weight Averaging-Gaussian(SWAG)

$$\mu_{SWA} = \frac{1}{T}\sum_{1}^{T}\theta_i$$

$$\Sigma_{SWA} = \frac{1}{T-1}\sum_{1}^{T}(\theta_i - \mu_{SWA})(\theta_i - \mu_{SWA})^T$$

dropout

$$p(y^*|x^*,X,Y) \approx \mathbb{E}_{\theta\sim q(\cdot|\lambda)}[p(y^*|x^*,\theta)]$$

dropout is also applied in prediction

- Bayesian Optimization

Mutual Information

$$F(s) = H(f) - H(f|y_s) = \frac{1}{2}log|I + \sigma^{-2}K_s|$$

$$F(S_T) \geq (1 - \frac{1}{e})\max_{S\subseteq D,|S|\leq T} F(S)$$

regret

$$R_T = \sum_{t=1}^{T}(max_{x\in D}f(x) - f(x_t))$$

- sublinear(optimal) if $\frac{R_T}{T} \to 0$
- $R_T^A \leq R_T^B$ cannot tell anything
- $\forall T\ R_T^A \leq R_T^B$ menas A is better

Bayesian Optimization

- Bayesian Optimization

Uncertainty Sampling

$$x_t = \underset{x\in D}{argmax}\,\sigma_{t-1}^2(x)$$

- maximizing information gain in homoscedastic noise case
- $\lim_{t\to\infty} \hat{x}_t = \underset{\in D}{argmax}\ \mu_t(x)$, $f(\hat{x}) \to f(x^*)$

Upper Confidence Sampling(UCB)

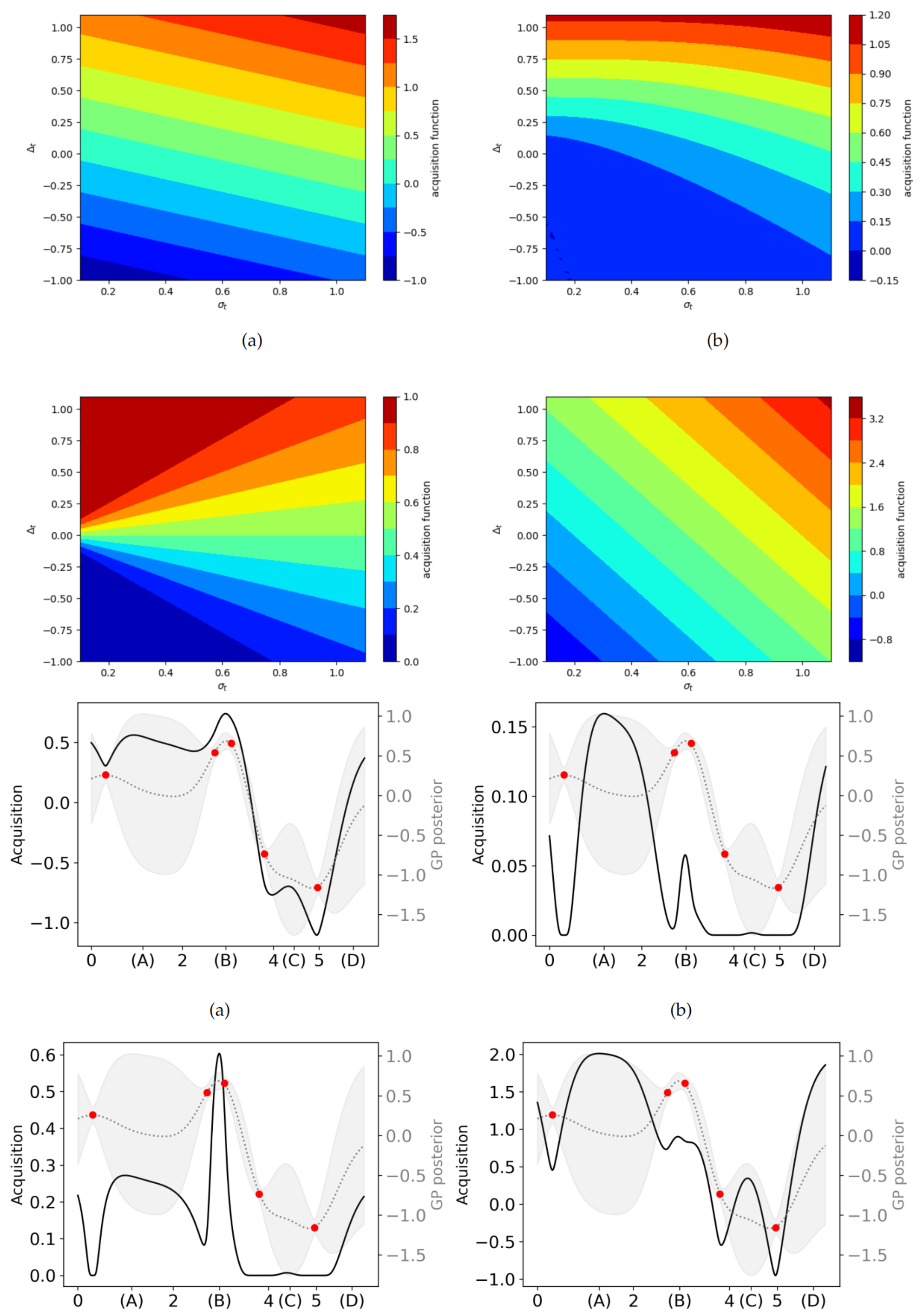
$$a = \mu_t(x) + \beta\sigma_t(x)$$

Probability of Improvement(PI)

$$a = \Phi\left(\frac{\mu_t(x) - f^*}{\sigma_t(x)}\right)$$

Expected Improvement(EI)

$$a = (\mu_t(x) - f^*)\Phi\left(\frac{\mu_t(x) - f^*}{\sigma_t(x)}\right) + \sigma_t(x)\phi\left(\frac{\mu_t(x) - f^*}{\sigma_t(x)}\right)$$



a:UCB($\beta = 0.5$) b:EI c:PI d:UCB($\beta = 2$)

$acquisition \uparrow \Longleftrightarrow exploitation$

$acquisition \downarrow \Longleftrightarrow exploration$

Reinforcement Learning

Bellman Theorem

$$V^*(x) = max_a(r(x,a) + \gamma\sum_{x'}P(x'|x,a)V^*(x'))$$

Hoeffding Bound

$$P(|\mu - \frac{1}{n}\sum_{i=1}^{n}Z_i| > \varepsilon) \leq 2exp(-\frac{2n\varepsilon^2}{C^2})$$

- Model Based

Value Iteration

- guarantee converge to an ε optimal policy not the exact optimal policy
- polynomial time, performance depend on the input

Policy Iteration

- guaranteed to monotonically improve the policy

ϵ greedy

when random number $< \epsilon$ do the random action

Rmax method

set reward R and transition probability $P(x^*|x,a) = 1$ at first

- with probability $1-\sigma$,reach ε - optimal
- polynomial time in $|X|, |A|, T, \frac{1}{\varepsilon}, log(\frac{1}{\delta})$

- Model Free

Temporal Difference(TD) - Learning

$$\hat{V}^\pi(x) \leftarrow (1-\alpha_t)\hat{V}^\pi(x) + \alpha_t(r + \gamma\hat{V}^\pi(x'))$$

on-policy

Theorem:$\sum_t\alpha_t = \infty, \sum_t\alpha_t^2 < \infty \Rightarrow P(\hat{V}^\pi \to V^\pi) = 1$

Q-Learning

off-policy

$$\hat{Q}^* \leftarrow (1-\alpha_t)\hat{Q}^*(x,a) + \alpha_t(r + \gamma\max_{a'}\hat{Q}^*(x',a'))$$

init : $\hat{Q}^*(x,a) = \frac{R_{max}}{1-\gamma}\Pi_{t=1}^{T_{init}}(1-\alpha_t)^{-1}$

Theorem: $\sum_t\alpha_t = \infty, \sum_t\alpha_t^2 < \infty \Rightarrow P(\hat{Q}^* \to Q^*) = 1$

- with probability $1-\sigma$, R max will reach an ε - optimal
- polynomial time in $|X|, |A|, T, \frac{1}{\varepsilon}, log(\frac{1}{\delta})$
- decay learning rate guarantee convergence

Deep RL

- Model Free

REINFORCE

$$\nabla_\theta J(\theta) = \mathbb{E}_{\tau\sim p_\theta(\tau)}[\sum_{t=0}^{T}\gamma^t(\sum_{t'=t}^{T}\gamma^{t'-t}r_t)\nabla_\theta ln\pi_\theta(a_t|x_t)]$$

$$\theta \leftarrow \theta + \eta_t\nabla_\theta J(\theta)$$

Actor-Critic

$$\nabla_{\theta_\pi}J(\theta_\pi) = \mathbb{E}_{(x,a)\sim\pi_{\theta_\pi}}[Q_{\theta_Q}(x,a)\nabla_{\theta_\pi}ln\pi_{\theta_\pi}(a|x)]$$

$$\theta_\pi \leftarrow \theta_\pi + \eta_t\nabla_{\theta_\pi}J(\theta_\pi)$$

$$\theta_Q \leftarrow \theta_Q - \eta_t(Q_{\theta_Q}(x,a) - r - \gamma Q_{\theta_Q}(x', \pi_{\theta_\pi}(x')))\nabla_{\theta_Q}Q_{\theta_Q}(x,a)$$

Proximal Policy Optimization(PPO)

$$L_{\theta_k}(\theta_k) = \mathbb{E}_{\tau\sim\pi_k}\sum_{t=0}^{\infty}\left[\frac{\pi_\theta(a|x)}{\pi_{\theta_k}(a|x)}(r + \gamma Q^{\pi_{\theta_k}}(x',a) - Q^{\pi_{\theta_k}}(x,a))\right]$$

$$\theta_k \leftarrow \theta_k - \eta_t\nabla_{\theta_k}L_{\theta_k}(\theta_k)$$

Deep Deterministic Policy Gradients(DDPG)

randomly add noise ensure sufficient exploration

on-policy	REINFORCE, Actor-Critic, PPO
off-policy	DDPG, TD3, SAC

- Model Based

reduce sample complexity

Random Shooting Method

Monte Carlo Tree Search

PETS