RV $\boldsymbol{A}\mathcal{N}_{\boldsymbol{\mu},\boldsymbol{\Sigma}} = \mathcal{N}_{\boldsymbol{A\mu},\,\boldsymbol{A\Sigma A}^\intercal}$ $\qquad |2\pi\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{\mu}\|^2_{\boldsymbol{\Sigma}^{-1}}}$ $\qquad (\boldsymbol{w}|\boldsymbol{v}) \sim \mathcal{N}_{\bar{\boldsymbol{w}}+\boldsymbol{\Sigma}_\times\boldsymbol{\Sigma}_{\boldsymbol{v}}^{-1}(\boldsymbol{v}-\bar{\boldsymbol{v}}),\,\boldsymbol{\Sigma}_{\boldsymbol{w}}-\boldsymbol{\Sigma}_\times\boldsymbol{\Sigma}_{\boldsymbol{v}}^{-1}\boldsymbol{\Sigma}_\times^\intercal}$

Proba $\underbrace{P_{X,Y} = P_{X|Y}P_Y}_{\text{chain/product rule}}$ $\quad \underbrace{P_{X=x} = \sum_y P_{X=x,Y=y}}_{\text{marginalization / sum rule}}$ $\quad \underbrace{\int p(y|f,X)p(f|\theta)df}_{p(y|X,\theta)\text{ model selection}}$ $\quad \underbrace{\mathbb{P}(|\mathbb{E}[f] - \frac{1}{N}\sum_{i=1}^N f(x_i)| > \varepsilon) \le 2e^{-\frac{2N\varepsilon^2}{\max f^2}}}_{\text{Hoeffding for a random sample } \{x_i\}_i}$

BL $\underbrace{p(\boldsymbol{y}|\boldsymbol{X},\theta) = \prod_i p(y_i|x_i,\theta)}_{\text{likelihood}}$ $\quad \underbrace{p(\theta)}_{\text{prior}}$ $\quad \underbrace{p(\theta|\boldsymbol{X},\boldsymbol{y}) \propto \text{likelihood} \times \text{prior}}_{\text{posterior}}$ $\quad \underbrace{p(y^*|x^*,\boldsymbol{X},\boldsymbol{y}) = \int p(y^*|x^*,\theta)\times[\text{posterior}]\,d\theta}_{\text{predictions}}$

Fact $\partial_\theta(\Sigma^{-1}) = -\Sigma^{-1}(\partial_\theta\Sigma)\Sigma^{-1}$ $\qquad \partial_\theta \ln|\Sigma| = (\Sigma^{-\intercal} : \partial_\theta\Sigma) = \operatorname{tr}(\Sigma^{-1}\partial_\theta\Sigma)$ $\qquad \underbrace{I(X,Y) = H[X] - H[X|Y]}_{\text{gain } \frac{1}{2}|I + \frac{1}{\sigma_\varepsilon^2}\mathbb{V}[X]| \text{ for } Y = X+\varepsilon}$

H $H[p,q] = -\mathbb{E}_p[\ln q]$ $\qquad (p\,\|\,q) = \mathbb{E}_p[\ln\frac{p}{q}]$ $\qquad H[\mathcal{N}] = \frac{1}{2}\ln((2\pi e)^d|\Sigma|)$ $\qquad (\mathcal{N}_0\,\|\,\mathcal{N}_1) = \frac{1}{2}(\operatorname{tr}(\frac{\boldsymbol{\Sigma}_0}{\boldsymbol{\Sigma}_1} - \boldsymbol{I}) + |\Delta\boldsymbol{\mu}|^2_{\boldsymbol{\Sigma}_1^{-1}} - \ln\frac{|\boldsymbol{\Sigma}_0|}{|\boldsymbol{\Sigma}_1|})$

BLR $y = X^\intercal w + \varepsilon$ $\qquad \mathbb{E}[w|y] = \bar{w} + K(y-\bar{y})$ $\qquad \mathbb{V}[w|y] = \boldsymbol{\Sigma}_w - \underbrace{\boldsymbol{\Sigma}_w X(\boldsymbol{\Sigma}_\varepsilon + X^\intercal\boldsymbol{\Sigma}_w X)^{-1}}_{\text{Kalman gain } K,\ \mathcal{O}(|y|^2)} X^\intercal\boldsymbol{\Sigma}_w \overset{\text{SMW}}{=} \underbrace{(\boldsymbol{\Sigma}_w^{-1} + X\boldsymbol{\Sigma}_\varepsilon^{-1}X^\intercal)^{-1}}_{\text{can be } \mathcal{O}(|w|^2\times|y|)}$

GP $\underbrace{w \overset{\text{prior}}{\sim} \mathcal{N}(\bar{w}, \Sigma_w \equiv k(\cdot,\cdot))}_{X := (\delta_x)_{x\in A}}$ $\quad \underbrace{\log p(x)}_{\text{evidence}} = \underbrace{(q\,\|\,p(\cdot|x))}_{\ge 0} + \underbrace{H[q] - H[q,p(\cdot,x)]}_{\text{ELBO}}$ $\quad \underbrace{\int (t-y)_+ p(y|x)dx}_{\text{EI}(x)\text{ over target } t}$ $\quad \underbrace{(Fw|y) + Bu' + \eta'}_{w_{\text{new}} \text{ in Kalman filter}}$

MC $\underbrace{\exists t : P^t > 0}_{\text{ergodic}} \Rightarrow \underbrace{\exists! \pi > 0}_{+X_\infty\sim\pi}$ $\quad \underbrace{P_{x'\leftarrow x}f_x = P_{x\leftarrow x'}f_{x'}}_{\text{det bal for } f \ \Rightarrow\ P*f=f}$ $\quad \underbrace{u \le \frac{R_{x\leftarrow x'}f_{x'}}{R_{x'\leftarrow x}f_x}}_{\text{MH } x'\leftarrow x}$ $\quad \underbrace{R^{\text{MALA}}_{\cdot\leftarrow x} = \mathcal{N}_{x-\tau g'_x,\,2\tau I}}_{g=-\log f:\ \cup\Rightarrow\text{convg}}$ $\quad \underbrace{e^{-(\lambda|\theta|^2 + \frac{n}{m}\sum_i^m \ln(1+e^{-y_i\theta\cdot x_i}))}}_{\text{log-reg } e^{-g}\quad \text{SGLD: } \not\nabla,\ \tau\sim t^{-\frac{1}{3}}}$

MDP $Q^V_{x,a} := r_{x,a} + \gamma\sum_{x'} V_{x'}P_{x'|x,a}$ $\quad \pi^Q_x := \arg Q_{x,\uparrow}$ $\quad \underbrace{V^\pi_x \overset{!}{=} Q^{V^\pi}_{x,\pi(x)}}_{\text{solve for } V^\pi}$ $\quad \underbrace{V^\star_x = Q^{V^\star}_{x,\uparrow}}_{\text{Bellman}}$ $\quad \underbrace{V^\pi \leftrightarrow \pi := \pi^{(Q^V)}}_{\text{policy it. convg}}$ $\quad \underbrace{V_x \leftarrow Q^V_{x,\uparrow}}_{\text{value it. } \varepsilon\text{-convg}}$

QL $\underbrace{Q_{s_t,a_t} \leftarrow (1-\alpha_t)\,Q_{s_t,a_t} + \alpha_t\,(r_t + \gamma Q_{s_{t+1},\uparrow})}_{(\text{RM for } \alpha)\,\wedge\,(\text{all } (s,a) \text{ recurrent}) \Rightarrow \text{convg}}$ $\quad \underbrace{\min_\theta \sum_{xarx'}[r + \gamma Q^{\theta^{\text{old}}}_{x',\uparrow} - Q^\theta_{x,a}]^2}_{\text{DQN}}$ $\quad \underbrace{\Sigma_t\varepsilon_t^2 < \infty \le \Sigma_t\varepsilon_t}_{\text{RM, } \epsilon\text{-greedy} \Rightarrow \text{convg}}$ $\quad \underbrace{r_{x,a} := R^{\max}}_{P?\ x^*!,\ \text{whp } \varepsilon\text{-cvg}}$

Policy∇ $\underbrace{\nabla_\theta\mathbb{E}[\Sigma_{t\ge1}\gamma^t r_t] = \mathbb{E}[\Sigma_{t\ge1}\Psi_t\nabla_\theta\ln\pi_\theta(a_t|s_t)]}_{\text{REINFORCE: for } t\ge1 \text{ do } \delta\theta = \eta\Psi_t\nabla_\theta\ln\pi_\theta}$ $\quad \Psi_t \in \{\Sigma(\gamma^{1:}r_{1:}),\ \underbrace{\Sigma(\gamma^{t:}(r_{t:} - b(x_{t:})))}_{\text{rew. to go } \Sigma(\gamma^{0:}r_{t:})},\ Q_t,\ \underbrace{Q_t - V_t}_{\text{adv.}},\ \underbrace{r_t + \gamma V_{t+1} - V_t}_{\text{TD: } \delta V = \alpha\times\text{TD}}\}$