Князев А, 491 группа

№1

$$E_{x,y}E_{X^l}(y-a_{X^l}(x))^2 = E_{x,y}E_{X^l}(y-E(y|x)+E(y|x)-a_{X^l}(x))^2 =$$

$$= \underbrace{E_{x,y}E_{X^l}[(y-E(y|x))^2]}_{\text{не зависит от } X^l} + E_{x,y}E_{X^l}[a_{X^l}-E_{X^l}a_{X^l}+E_{X^l}a_{X^l}-$$

$$-E(y|x)]^2 + \underbrace{2E_{x,y}E_{X^l}[(y-E(y|x))\cdot(E(y|x)-a_{X^l}(x))]}_{a} \ominus$$

$$a = 2\underbrace{E_{x,y}(y-E(y|x))}_{=0}E_{X^l}[E(y|x)-a_{X^l}(x)] = 0$$

$$\ominus\ E_{x,y}[(y-E(y|x))^2] + E_{x,y}E_{X^l}[\overbrace{E_{X^l}a_{X^l}-E(y|x)}^{\text{не зависит от } X^l}]^2 +$$

$$+E_{x,y}E_{X^l}[a_{X^l}-E_{X^l}a_{X^l}]^2 + 2E_{x,y}E_{X^l}[[a_{X^l}-E_{X^l}a_{X^l}]\cdot[E_{X^l}a_{X^l}-E(y|x)]] =$$

$$= noise + bias + variance +$$

$$+2E_{x,y}[[E_{X^l}a_{X^l}-E(y|x)]\underbrace{E_{X^l}[-E_{X^l}a_{X^l}+a_{X^l}]}_{=0}] =$$

$$= noise + bias + variance$$

■

№2

$$a(x) = \frac{1}{N}\sum_{n=1}^{N} a_n(x)$$

опущен индекс алгоритма

Смещение: $E_{x,y}\left[\left[E_{X^e}\left(\frac{1}{N}\sum_{n=1}^{N} a_{X^e}(x)\right) - E(y|x)\right]^2 \right] =$

$= E_{x,y}\left[\left[\frac{1}{N}\sum_{n=1}^{N} E_{X^e}\left[a_{X^e}(x) - E(y|x)\right]\right]^2 \right] =$

$= E_{x,y}\left[\left(E_{X^e}\left[a_{X^e}(x) - E(y|x)\right]\right)^2 \right] =$

$= E_{x,y}\left[\left[E_{X^e}\left[a_{X^e}(x)\right] - E(y|x)\right]^2 \right]$

смещение композиции, полученной с помощью бэггинга, совпадает со смещением одного алгоритма

Разброс: $E_{x,y}\left[E_{X^e}\underbrace{\left[\left(\frac{1}{N}\sum_{n=1}^{N} a_{X^e}(x) - E_{X^e}\left[\frac{1}{N}\sum_{n=1}^{N} a_{X^e}(x)\right]\right)^2 \right]}_{(1)} \right]$

$(1) = \frac{1}{N^2}\left(\sum_{n=1}^{N} \left[a_{X^e}(x) - E_{X^e}\left[a_{X^e}(x)\right]\right]\right)^2 =$

$= \frac{1}{N^2}\sum_{n=1}^{N}\left(a_{X^e} - E_{X^e}\left[a_{X^e}(x)\right]\right)^2 +$

$+ \frac{1}{N^2}\sum_{n\neq k}\left[\left[a_{X^e} - E_{X^e}\left[a_{X^e}(x)\right]\right]\cdot\left[a_{X^e}(x) - E_{X^e}\left[a_{X^e}(x)\right]\right]\right]$

для разных алгоритмов

$E_{x,y}E_{X^e}\left[(1)\right] = \frac{1}{N^2}E_{x,y}\left[E_{X^e}\left[\sum_{n\neq k}(a_{X^e} - Ea_{X^e})\cdot(a_{X^e} - Ea_{X^e})\right]\right] +$

$+ \frac{1}{N^2}E_{x,y}\left[E_{X^e}\left[\sum_{n=1}^{N}(a_{X^e} - Ea_{X^e})^2\right]\right] =$

$= \frac{1}{N}E_{x,y}\left[E_{X^e}\left[\left(a_{X^e}(x) - E_{X^e}\left[a_{X^e}(x)\right]\right)^2\right]\right] +$

$$+\frac{N(N-1)}{N^2} E_{x,y}\left[E_{X^\ell}\left([a_{X^\ell}-E_{X^\ell} a_{X^\ell}]\cdot[a_{X^\ell}-E_{X^\ell} a_{X^\ell}]\right)\right]=$$

для любых двух различных алгоритмов

$$= a + b$$

a – ~~дисперс~~ разброс базового алгоритма, деленный на длину композиции N

b – ковариация между двумя алгоритмами

Если алгоритмы некоррелированы, то разброс в N раз меньше разброса одного алгоритма! ~~[illegible]~~

№3

$$\operatorname{corr}(\varepsilon;\eta)=\frac{\operatorname{cov}(\varepsilon;\eta)}{\sigma_\varepsilon\cdot\sigma_\eta}$$

$D(x_i;x_j)=\sigma^2;\quad \operatorname{corr}(x_i;x_j)=\rho,$

$$D\left[\frac{x_1+\dots+x_m}{m}\right]=E\left[(\dots)^2\right]-\left[E(\dots)\right]^2=$$

$$=\frac{1}{m^2}\left[\sum_{i=1}^{m}\left[Ex_1^2\right]+m(m-1)E(x_i\cdot x_j)\right]-\frac{1}{m}\cdot m\cdot\left[Ex_1\right]^2=$$

$$=\frac{1}{m^2}\left[m\cdot\left(\sigma^2+(Ex_1)^2\right)+m(m-1)\left[\sigma^2\cdot\rho+(Ex_1)^2\right]\right]-$$

$$-(Ex_1)^2=\frac{\sigma^2}{m}+\cancel{\frac{(Ex_1)^2}{m}}+\sigma^2\rho+\cancel{(Ex_1)^2}-\frac{\sigma^2\rho+\cancel{(Ex_1)^2}}{m}-$$

$$-\cancel{(Ex_1)^2}=\rho\sigma^2+(1-\rho)\frac{\sigma^2}{m}$$