2.1. Bias-variance-noise decomposition

Д-ть $E_{x,y} E_{X^\ell} (y - a_{X^\ell}(x))^2 =$

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

$+ E_{x,y} E_{X^\ell} (a_{X^\ell}(x) - E_{X^\ell} a_{X^\ell}(x))^2$

Δ. 1) распишем ~~пр~~ левую часть:

$E_{x,y} E_{X^\ell} y^2 - 2 E_{x,y} E_{X^\ell} y a_{X^\ell}(x) + E_{x,y} E_{X^\ell} (a_{X^\ell}(x))^2$

2) правую часть:

noise: $E_{x,y} y^2 - 2E_{x,y}(y E(y|x)) + E_{x,y}(E(y|x))^2 +$

bias: $+ E_{x,y}(E(y|x))^2 - 2E_{x,y}(E(y|x)) E_{X^\ell} a_{X^\ell}(x)) +$

$+ E_{x,y}(E_{X^\ell} a_{X^\ell}(x))^2 +$

var: $+ E_{x,y} E_{X^\ell} (a_{X^\ell}(x))^2 - 2E_{x,y} E_{X^\ell} (a_{X^\ell}(x) \cdot E_{X^\ell} a_{X^\ell}(x)) +$

$+ E_{x,y} E_{X^\ell} (E_{X^\ell} a_{X^\ell}(x))^2$

3) Подчёркнутые ~~~ слагаемые совпадают ⇒ можем их сократить

Подчёркнутые - - - слагаемые складываем

Подчёркнутые ⌇⌇⌇ равны: $E_{x,y} E_{X^\ell} y^2 = E_{x,y} y^2$

(т.к. при данном x для любых ответов алгоритма y постоянная величина и $E_{X^\ell} y = y$).

Также можем преобразовать в левой части

$-2E_{x,y} E_{X^\ell} y a_{X^\ell}(x) = -2E_{x,y}\, y\, a_{X^\ell}(x)$

4) Получим равенство:

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

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

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

$$-2E_{x,y}\, E_{X^l}(a_{X^l}(x) \cdot E_{X^l}\, a_{X^l}(x)) +$$

$$+E_{x,y}\, E_{X^l}(E_{X^l}\, a_{X^l}(x))^2$$

Заметим, что $+2E_{x,y}(y\, E(y|x)) \overset{\text{п.н.}}{=} +2E_{x,y}(E(y|x))^2$

т.к. $y \overset{\text{п.н.}}{=} E(y|x)$ ~~вуп~~

5) Получим:

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

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

$$+E_{x,y}\, E_{X^l}(E_{X^l}\, a_{X^l}(x))^2$$

Заметим, что $E_{x,y}(E_{X^l}\, a_{X^l}(x))^2 = E_{x,y}\, E_{X^l}(E_{X^l}\, a_{X^l}(x))^2$

6) Получим:

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

распишем левую часть: $E_{x,y}\, y\, a_{X^l}(x) = E_{x,y}\, y\, E_{X^l} a_{X^l}(x) =$

$$= E_{x,y}(E(y|x)\, E_{X^l}\, a_{X^l}(x)) = \cancel{E_{x,y}}$$

7) тождество $\Rightarrow$ начальное выр-е тоже тождество. □

2.2 Смещение и разброс в бэггинге

$$a(x) = \frac{1}{M}\sum_{m=1}^{M} a_m(x)$$

Δ. 1) $bias(a(x)) = E_{x,y}\left(E(y|x) - E_{X^\ell}\left(\frac{1}{M}\sum_{1}^{M} a_m(x)\right)\right)^2 =$

$= E_{x,y}\left(E(y|x) - \frac{1}{M}\sum_{1}^{M} E_{X^\ell} a_m(x)\right)^2 =$ / ответы всех базовых алгоритмов распределены одинаково /

$= E_{x,y}\left(E(y|x) - E_{X^\ell} \overset{\forall i = \overline{1,M}}{a_i(x)}\right) = bias(a_i(x))$

2) $var(a(x)) = E_{x,y} E_{X^\ell}\left(\frac{1}{M}\sum(a_i(x)) - E_{X^\ell}(a_i(x))\right)^2 =$

$= \frac{1}{M^2} E_{x,y} E_{X^\ell}\left(\sum(a_i(x) - E_{X^\ell}(a_i(x)))\right)^2 =$

/ a_i некоррелируемы при $i \neq j$: $\underbrace{E_{X^\ell}(a_i(x) - E_{X^\ell}(a_i(x)))(a_j(x) - E_{X^\ell}(a_j(x)))}_{=0}$ /

$= \frac{1}{M^2} E_{x,y} E_{X^\ell} M(a_i(x) - E_{X^\ell}(a_i(x)))^2 =$

$= \frac{1}{M} \underbrace{E_{x,y} E_{X^\ell}(a_i(x) - E_{X^\ell}(a_i(x)))^2}_{Var(a_i(x)),\ \forall i = \overline{1,M}} =$

$= \frac{1}{M} var(a_i(x))$ $\square$.

<u>Ответ</u>: $bias(a(x)) = bias(a_i(x)) \quad \forall i$

$var(a(x)) = \frac{1}{M} var(a_i(x)) \quad \forall i$

2.3 Корреляция ответов базовых алгоритмов

n одинаково распределённых случ. величин с дисперсией σ^2. $\forall i, j;\ i \neq j \mapsto \operatorname{corr}(X_i, X_j) = \rho$

д-ть, что $D\left(\frac{1}{n}\sum X_i\right) = \rho\sigma^2 + (1-\rho)\frac{\sigma^2}{n}$.

$$\Delta.\quad D\left(\frac{1}{n}\sum X_i\right) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \underbrace{E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)}_{EX_1}\right)^2 =$$

$$= \frac{1}{n^2} E\left(\sum_{1}^{n}(X_i - EX_1)\right)^2 =$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} E(X_i - EX_1)^2 + \sum_{i}^{n}\sum_{j,\, i\neq j}^{n}(X_i - EX_1)(X_j - EX_1)\right) =$$

$$= \frac{1}{n}\sigma^2 + 2\underbrace{C_n^2}_{\frac{n!}{2(n-2)!}}\rho\sigma^2 = \frac{1}{n}\sigma^2 + \rho\sigma^2 - \rho\frac{\sigma^2}{n} =$$

$$= \rho\sigma^2 + (1-\rho)\frac{\sigma^2}{n}$$

$\square$.