(2.1) Будем считать, что на $X \times Y$ сущ-ет распределение $p(x,y)$.

$$E_{X^\ell} E_{xy} (y - a(x))^2 - ?$$

$$E_{xy}(y-a(x))^2 = E_{xy}(y - E(y|x) + E(y|x) -$$

$$- a(x))^2 = E_{xy}(y - E(y|x))^2 + 2E_{xy}((y - E(y|x$$

$$\cdot (E(y|x) - a(x))) + E_{xy}(E(y|x) - a(x))^2$$

$$2E_{xy}(y - E(y|x))(E(y|x) - a(x)) = 2\int_Y \int_X (y -$$

$$- E(t|x)) \cdot (E(t|x) - a(x)) p(x,y)\, dx\, dy =$$

$= |[E(t|x) - a(x)]$ нез от $y \Rightarrow$ вынесем$| =$

$$= \int_X (E(t|x) - a(x)) \cdot \int_Y (y - E(t|x)) p(x,y) dx\, dy =$$

$$= \int_X (E(t|x) - a(x)) \cdot \Big( \int_Y y p(x,y) dy - \int_Y E(t|x) p(x,y) dy$$

$$p(x,y) = p(x) p(y|x) \Big| = \int_X (E(t|x) - a(x)) \cdot$$

$$\Big( p(x) \int_Y y p(y|x) - E(t|x) \int_Y p(x|y) dy \Big) dx =$$

$$= \int_X (E(t|x) - a(x))(\underbrace{p(x)E(t|x) - p(x)E(t,x)}_{=0})\,dx$$

$$E_{xy}(y-a(x))^2 = E_{xy}(y-E(y|x))^2 + E_{xy}(E(y|x)-a(x))^2$$

ш $X^e$

$$E_{X^e}E_{xy}(y-a(x))^2 = E_{X^e}E_{xy}(y-E(y|x))^2 +$$

$$+ E_{X^e}E_{xy}(E(y|x)-a(x))^2 = E_{xy}(y-E(y|x))^2 +$$

$$+ E_{xy}(E_{X^e}(E(y|x)-a(x))^2)$$

$$E_{xy}E_{X^e}(E(y|x)-a(x))^2 = E_{xy}E_{X^e}(E(y|x) -$$

$$- E_{X^e}a(x) + E_{X^e}(a(x)) - a(x))^2 = |ш X^e| =$$

$$= E_{xy}E_{X^e}(E(y|x) - E_{X^e}(a(x)))^2 + E_{xy}E_{X^e}($$

$$E_{X^e}a(x) - a(x))^2 + 2E_{xy}E_{X^e}(E(y|x) -$$

$$- E_{X^e}(a(x)))(E_{X^e}a(x) - a(x))$$

ита

$$E_{X^e}(E(y|x) - E_{X^e}a(x))(E_{X^e}a(x) - a(x)) =$$

ш $X^e$

$$= (E(y|x) - E_{X^e}a(x))E_{X^e}(E_{X^e}a(x) - a(x)) =$$

$$= (E(y|x) - E_{X^e}a(x))(\underbrace{E_{X^e}(a(x)) - E_{X^e}(a(x))}_{=0})$$

$\Rightarrow$

$$\Rightarrow E_{X^l} E_{xy}(y-a(x))^2 = \underbrace{E_{xy}(y-E(y/x))^2}_{\text{шум}} +$$

$$+ \underbrace{E_{xy}(E_{X^l}a(x)-E(y/x))^2}_{\text{смещ.}} + \underbrace{E_{xy}E_{X^l}(a(x)-E_{X^l}(a(x)))^2}_{\text{разброс}}$$

2.2 $a(x) = \frac{1}{M}\sum_{m=1}^{M} a_m(x)$

Обуч алгорит $a_m(x)$ с пом метода $\tilde{\mu}$ и ответ $a(x) = \frac{1}{N}\sum_{m=1}^{N} a_m(x) = \frac{1}{N}\sum_{m=1}^{N}\tilde{\mu}(X^l)(x)$

$$E_{xy}\left(E_{X^l}\left(\frac{1}{N}\sum_{n=1}^{N}\tilde{\mu}(X^l)(x)\right) - E(y/x)\right)^2 =$$

$$= E_{xy}\left(\frac{1}{N}\sum E_{X^l}\tilde{\mu}(X^l)(x) - E(y/x)\right)^2 =$$

$$= E_{xy}\left(E_{X^l}\tilde{\mu}(X^l)(x) - E(y/x)\right)^2$$

смещ композ = смещ. одного алгор.

$$E_{xy}E_{X^l}\left(\frac{1}{N}\sum\tilde{\mu}(X^l)(x) - E_{X^l}\left(\frac{1}{N}\sum\tilde{\mu}(X^l)(x)\right)\right)^2 +$$

$$\frac{1}{N}\sum\tilde{\mu}(X^l)(x) - E_{X^l}\left(\frac{1}{N}\sum\tilde{\mu}(X^l)(x)\right) = \frac{1}{N^2}\cdot$$

$$\cdot\left(\sum\left(\tilde{\mu}(X^l)(x) - E_{X^l}\tilde{\mu}(X^l)(x)\right)\right)^2 = \frac{1}{N^2}\sum\left(\tilde{\mu}(X^l)(x) -\right.$$

$$\left.- E_{X^l}\tilde{\mu}(X^l)(x)\right)^2 + \frac{1}{N^2}\sum_{n_1\neq n_2}\left(\tilde{\mu}(X^l)(x) - E_{X^l}\tilde{\mu}(X^l)(x)\right)\cdot$$

$\cdot(\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x)) = A$

$E_{xy}E_{X^\ell}(A) = \frac{1}{N^2}E_{xy}E_{X^\ell}(\sum(\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x)$

$+\frac{1}{N^2}E_{xy}E_{X^\ell}(\sum_{n_1\neq n_2}(\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x)) \cdot$

$\cdot(\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x)) = \frac{1}{N}E_{xy}E_{X^\ell}($

$\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x))^2 + \frac{N(N-1)}{2}E_{xy}E_{X^\ell}($

$(\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x))\cdot(\tilde{\mu}(X^\ell)(x) - E_{X^\ell}\tilde{\mu}(X^\ell)(x))$

1ое слаг – $\sigma^2$ одного алг. (базов)
2ое слаг – ков между двумя баз

Если ~~д~~ алгор не коррелир => дисперсия уменьш. в n раз

(2.3) $a_i$ – од распр

$$a(x) = \frac{1}{k}\sum a_i(x)$$

$E_{x,y,x,y}(a(x)) = \frac{1}{k}E_{xy_{xy}}(a_i(x)) = E_{xy_{xy}}(a_1(x)) \Rightarrow$

=> bias не уменьшится

$Var_{xy_{xy}}(a(x))^2 = \frac{1}{k^2}Var_{xy_{xy}}(\sum_{ij} a_i(x)a_j(x))$

$= \frac{1}{k}Var_{xy_{xy}}\, a_1(x) + \frac{1}{k^2}\sum_{i\neq j} cov(a_i(x), a_j(x))$

Если коэф корр один. у всех пар ($= r$)

$$\frac{1}{k} Var_{XY_x} a_1(x) + \frac{1}{k^2} \sum_{i \neq j} cov(a_i(x), a_j(x)) =$$

$$= \frac{1}{k} Var_{XY_x} a_1(x) + \frac{1}{k^2} \sum_{i \neq j} r \, Var_{XY_x} a_1(x) =$$

$$= \left( \frac{1}{k} + \frac{r(k-1)}{k} \right) Var_{XY_x} a_1(x)$$

$\Rightarrow$ Var меньше, чем меньше корр.