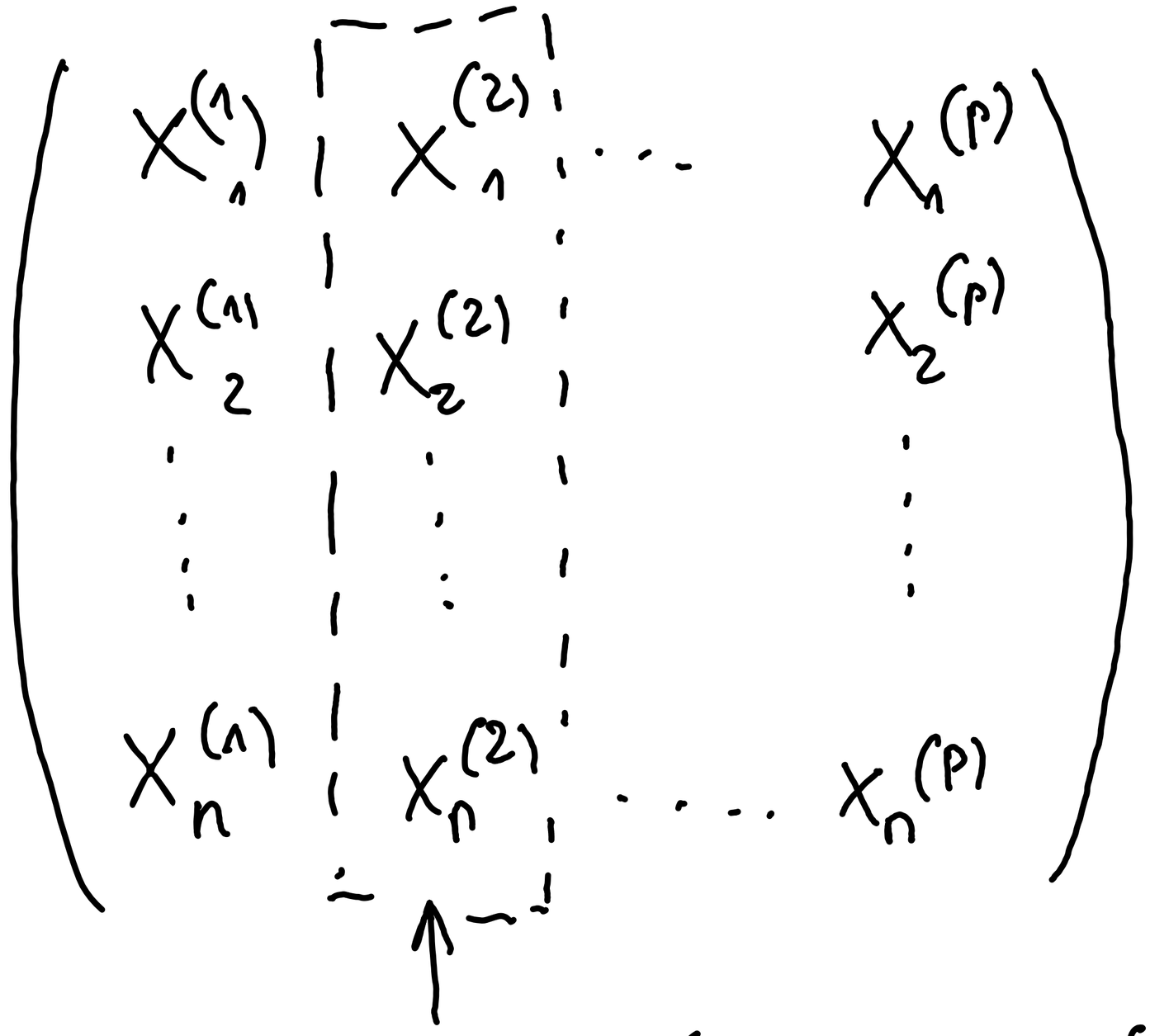
Variable importance using randomization



Importance of variable $X^{(j)}$:

- MSE from OOB predictions: $MSE^{(OOB)}$
- MSE from OOB predictions, but including a permuted version of $X^{(j)}$ (leaving all other variables unchanged) $MSE^{(OOB)}(X^{(j)})$

$MSE^{(OOB)}(x^{(j)}) - MSE^{(OOB)} = \Delta MSE(j)$

is the increase in MSE when $x^{(j)}$ is permuted.

$\Rightarrow$ If this is large, then $x^{(j)}$ is an important predictor.

Note: For every datapoint we use the predictions from the $B/3$ trees that were built without using that datapoint in the bootstrap sample.