

Exercise 6

Philipp Hanslovsky, Robert Walecki

June 10, 2012

6.3.1

The covariance is a symmetric bilinear form. Therefore the following rules apply:

$$Cov(X, Y) = Cov(Y, X) \quad (1)$$

$$Cov(aX + b, Y) = a * Cov(X, Y) \quad (2)$$

$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z) \quad (3)$$

3 can be generalized to:

$$Cov(\sum_i X_i, Y) = \sum_i Cov(X_i, Y) \quad (4)$$

Furthermore the covariance is a generalization of the variance:

$$Var(X) = Cov(X, X) \quad (5)$$

The relation between covariance and correlation is given by:

$$\rho_{ij} = \frac{cov(X_i, X_j)}{\sigma_i \sigma_j} \stackrel{i.d.}{=} \frac{cov(X_i, X_j)}{\sigma^2} \quad (6)$$

With this preconditions we get:

$$Var(\frac{1}{B} \sum_{i=1}^B X_i) \stackrel{5}{=} Cov(\frac{1}{B} \sum_{i=1}^B X_i, \frac{1}{B} \sum_{i=1}^B X_i) \quad (7)$$

$$\stackrel{2,3}{=} \frac{1}{B^2} \sum_{i,j=1}^B (X_i, X_j) \quad (8)$$

$$\stackrel{5}{=} \frac{1}{B^2} \sum_{i=1}^B Var(X_i) + \frac{2}{B^2} \sum_{i=1}^{B-1} \sum_{j=i+1}^B cov(X_i, X_j) \quad (9)$$

$$\stackrel{6,i.d.}{=} \frac{1}{B^2} B \sigma^2 + \frac{2\rho\sigma^2}{B^2} \sum_{i=1}^{B-1} \sum_{j=i+1}^B 1 \quad (10)$$

$$= \frac{\sigma^2}{B} + \frac{2\rho\sigma^2}{B^2} \sum_{i=1}^{B-1} (B - i) \quad (11)$$

$$= \frac{\sigma^2}{B} + \frac{2\rho\sigma^2}{B^2} \left(B(B-1) - \frac{B(B-1)}{2} \right) \quad (12)$$

$$= \frac{\sigma^2}{B} + \frac{\rho\sigma^2}{B} (B-1) \quad (13)$$

$$= \rho\sigma^2 + \frac{(1-\rho)\sigma^2}{B} \quad (14)$$

6.3.2

The probability p_{oob} for an observation to be out of bag is given by the ratio of the number N_{all} of all possible bags and the number N_{-1} of all possible bags that do not contain given observation. Let N be the number of all observations and k the number of elements contained in a bag. Then the total number of possible bags is given by the number of possible configurations $\frac{N!}{(N-k)!}$ divided by the number of corresponding permutations $k!$ (i.e. choosing observations 1, 2 and 4 is the same as choosing 1, 4 and 2) or the binomial coefficient.

$$N_{all} = \frac{N!}{k!(N-k)!} \quad (15)$$

$$= \binom{N}{k} \quad (16)$$

The number of possible bags not containing a certain observation can be obtained in a similar fashion. Let x_i be an observation fixed to be not in a bag. Then the number of possible bags for that configuration is given by the number of possible configurations for the remaining $N-1$ samples $\frac{(N-1)!}{(N-k-1)!}$ divided by the number of corresponding permutations $k!$.

$$N_{-1} = \frac{(N-1)!}{k!(N-1-k)!} \quad (17)$$

$$= \binom{N-1}{k} \quad (18)$$

Therefore the resulting probability p_{oob} is given by:

$$p_{oob} = \frac{N_{-1}}{N_{all}} \quad (19)$$

$$= \frac{(N-1)!}{k!(N-1-k)!} \frac{k!(N-k)!}{N!} \quad (20)$$

$$p_{oob} = \frac{N-k}{N} = 1 - \frac{k}{N} \quad (21)$$