Proof: Follows directly from the symmetry of covariance.

(2) (Shift-invariance) Suppose X, Y are jointly distributed nondegenerate r.v.s having finite 2^{nd} moments. Then for any $b, d \in \mathbb{R}$,

$$\rho(X+b, Y+d) = \rho(X,Y).$$

Proof: Follows from the shift-invariance of covariance and variance:

$$\left.\begin{aligned} \mathrm{Cov}(X+b, Y+d) &= \mathrm{Cov}(X,Y) \\ \mathrm{Var}(X+b) &= \mathrm{Var}(X) \\ \mathrm{Var}(Y+d) &= \mathrm{Var}(Y) \end{aligned}\right\}$$

Exc: Prove these.

(3) (Scale-invariance of the absolute value of correlation)

Suppose X, Y are jointly distributed nondegenerate r.v.s having finite 2^{nd} moments. Then for any $a, c \in \mathbb{R}$, the following trichotomy holds:

$$\rho(aX, cY) = \begin{cases} \rho(X,Y) & \text{if } ac > 0, \\ \text{undefined} & \text{if } ac = 0, \\ -\rho(X,Y) & \text{if } ac < 0. \end{cases}$$

(219)

Proof: If $ac = 0$, then either $a = 0$ or $c = 0$, which yields $\left(Var(aX) = 0 \text{ or } Var(aY) = 0 \right)$ and $Cov(aX, cY) = ac\, Cov(X, Y) = 0$ leading to a $\frac{0}{0}$ form for ~~corre~~ $\rho(aX, cY)$.

Assume now that $ac \neq 0$. Then

$$\rho(aX, cY) = \frac{Cov(aX, cY)}{+\sqrt{Var(aX)\, Var(cY)}}$$

$$= \frac{ac\, Cov(X, Y)}{+\sqrt{a^2\, Var(X)\, c^2\, Var(Y)}}$$

$$= \frac{ac}{|ac|} \frac{Cov(X, Y)}{+\sqrt{Var(X)\, Var(Y)}}$$

$$= \frac{ac}{|ac|} \rho(X, Y)$$

$$= \begin{cases} \rho(X, Y) & \text{if } ac > 0, \\ -\rho(X, Y) & \text{if } ac < 0. \end{cases}$$

(4) (Correlation is invariant under change of units)

Suppose X and Y are jointly distributed nondegenerate r.v.s having finite 2^{nd} moments. Take $a, b, c, d \in \mathbb{R}$ such that $a \neq 0$ and $c \neq 0$. Then the following dichotomy holds:

$$\rho(aX+b, \cancel{cX+d}\ cY+d) = \begin{cases} \rho(X,Y) & \text{if } ac>0, \\ -\rho(X,Y) & \text{if } ac<0. \end{cases}$$

Proof: $(2)+(3) \Rightarrow (4)$.

Remarks: ① Suppose we want to measure the correlation coefficient between the daily maximum temperature of Bangalore in the months of January and March. Property (4) says that the value of ~~the~~ this correlation coefficient remains unchanged even if we change the unit of temperature from °C to ~~[illegible]~~ °F.

② So far, we have only discussed properties of correlation coefficient that makes it a good measure of association. Now we shall discuss a property that tells us that correlation essentially measures (the amount and the direction of) the linear association. The proof of the following property follows from the Cauchy-Schwarz Inequality (see Pg 210) and the discussions in Pg 213 - 216.

(5) Suppose X and Y are jointly distributed nondegenerate r.v.s having finite 2^{nd} moment.

Then $-1 \leq \rho(X,Y) \leq +1$.

Moreover,

(a) $\rho(X,Y) = 1$ holds if and only if there exist $a > 0$ and $b \in \mathbb{R}$ such that

$P[Y = aX + b] = 1$. (Complete (or perfect) positive linear association)

(b) $\rho(X,Y) = -1$ holds if and only if there exist $a < 0$ and $b \in \mathbb{R}$ such that

$P[Y = aX + b] = 1$. (Complete (or perfect) negative linear association)

(c) $X \perp\!\!\!\perp Y \Rightarrow \rho(X,Y) = 0$ but

$\rho(X,Y) = 0 \not\Rightarrow X \perp\!\!\!\perp Y$.
(does not always imply)

Proof: Let $\mu_x = E(X)$ and $\mu_y = E(Y)$. Both of these exist and are finite because of the hypothesis that X, Y have finite 2^{nd} moments.

Applying Cauchy-Schwarz inequality (see Pg 210) on the jointly distributed r.v.s $U = X - \mu_x$ and $V = Y - \mu_Y$ with finite 2^{nd} moments, we get

$$|E(UV)| \leqslant +\sqrt{E(U^2)\,E(V^2)}$$

$$\Leftrightarrow \left|E\left[(X-\mu_x)(Y-\mu_y)\right]\right| \leqslant +\sqrt{E\left[(X-\mu_x)^2\right]\,E\left[(Y-\mu_y)^2\right]}$$

$$\Leftrightarrow |\mathrm{Cov}(X,Y)| \leqslant +\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}$$

$$\Leftrightarrow \frac{|\mathrm{Cov}(X,Y)|}{+\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \leq 1 \qquad \left[\because X, Y \text{ nondegenerate} \Rightarrow \mathrm{Var}(X) > 0 \text{ and } \mathrm{Var}(Y) > 0\right]$$

$$\Leftrightarrow |\rho(X,Y)| \leq 1$$

$$\Leftrightarrow -1 \leqslant \rho(X,Y) \leq 1.$$

This completes the proof of the inequalities in Property (5).

(a) Only if part

~~Following the strin~~ Chasing back the steps in Pg 222, we get that $\rho(X, Y) = 1$ holds if and only if $E(UV) = +\sqrt{E(U^2)E(V^2)}$ holds, where $U = X - \mu_X$ and $V = Y - \mu_Y$.

Therefore, using the conclusion given at the beginning of Pg 216, it follows that

$\rho(X, Y) = 1 \implies$ either $P[U = 0] = 1$

or $\exists\ \gamma \in [0, \infty)$ such that $P[V = \gamma U] = 1$.

Since X is assumed to be nondegenerate, we get $P[U=0] = P[X = \mu_X] < 1$. Therefore $\rho(X,Y) = 1 \Rightarrow \exists\ \gamma \in [0, \infty)$ such that $P[V = \gamma U] = 1$. Again, since Y is assumed to be nondegenerate, we get $\gamma \neq 0$. This is because $\gamma = 0 \Rightarrow P[V = 0] = P[Y = \mu_Y] = 1$.

Hence $\rho(X,Y) = 1 \Rightarrow \exists\ \gamma \in (0, \infty)$ such that

$$P[V = \gamma U] = 1$$

$$\Leftrightarrow P[Y - \mu_Y = \gamma(X - \mu_X)] = 1$$

$$\Leftrightarrow P[Y = \gamma X + (\mu_Y - \gamma \mu_X)] = 1$$

$\Leftrightarrow P[Y = aX + b] = 1$, where $a = \gamma \in (0, \infty)$

and $b = \mu_Y - \gamma \mu_X \in \mathbb{R}$.

This concludes the proof of "only if part".

<u>If part</u>

Suppose $\exists$ $a \in (0, \infty)$ and $b \in \mathbb{R}$ such that $P[Y = aX + b] = 1$. We have to show that $\rho(X, Y) = 1$. Note that the event $(Y = aX + b)$ implies the event $\cancel{[(X, Y) = (X, aX+b}$ $[(X, Y) = (X, aX + b)]$. In other words,

$$(Y = aX + b) \subseteq [(X, Y) = (X, aX + b)]$$

leading to

$$1 = P[Y = aX+b] \leq P[(X,Y) = (X, aX+b)],$$

which yields $P[(X,Y) = (X, aX+b)] = 1$.

Clearly, this means that the random vectors (X,Y) and $(X, aX+b)$ are equal with probability 1 and hence they must have the same joint distribution. This implies, in particular, that

$$\rho(X,Y) = \rho(X, aX+b)$$

$$= \rho(X,X) \quad \text{[Using } a>0 \text{ and Property (4) in Pg 220]}$$

$$= 1 \quad \text{[Using Property (0) of Pg 217]}$$

completing the proof of "if part".

(b) The proof is left as an exercise.

Exc: Prove Property (5) (b).

(c) Follows from the corollary stated at the end of Pg 186 + Remark 1 of Pg 184.

Remarks: 1 Please revisit Remarks 1 - 4 of Pg 184 - 185 in light of ~~the~~ Property (5).

2 It is now clear that correlation coefficient ~~measures~~ is a measure of linear association. Its sign gives the direction (positive ~~of~~ or negative) of linear association and its absolute value gives the amount of linear association between two jointly distributed ~~nondeg~~ nondegenerate r.v.s with finite second moments.

Exc: Compute $\rho(X,Y)$ for each of the random vectors (X,Y) described/discussed in the examples given in the pages mentioned below, and heuristically justify its sign:
(i) Pg 13, (ii) Pg 18 with $p=\frac{1}{2}$, (iii) Pg 31.