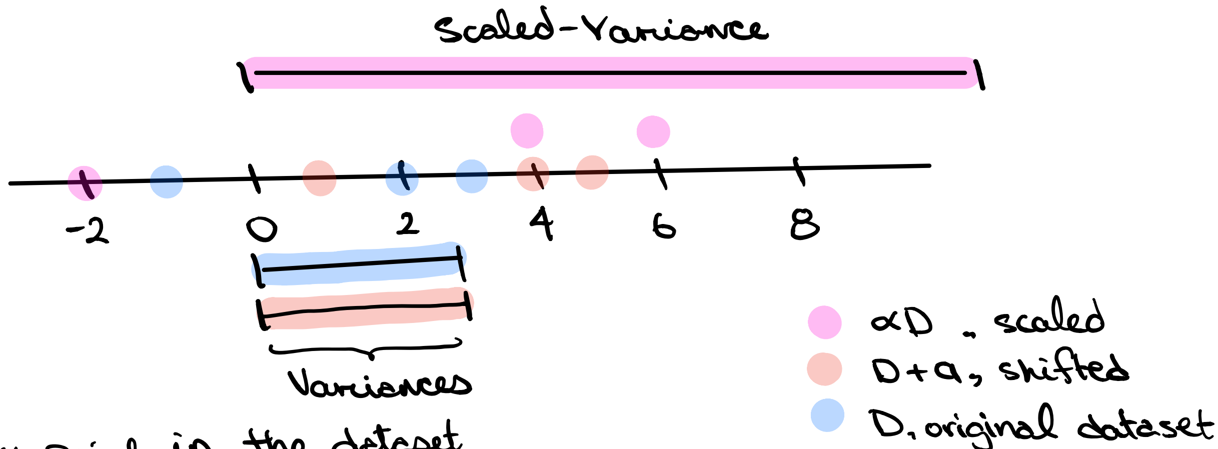
Linear transformations performed on a dataset will change both the data and it's corresponding description – i.e., the summary statistics describing that data.

This is true for co/variance.



- If every point in the dataset is shifted, the mean is also shifted. Thus the deviations from every point to the mean is the same since both were translate with 1:1 covariance.

- If every point is scaled, the mean will likewise be scaled. The distances between the points and mean will by nature scale equally; this means the that the squared deviations will be increased by a factor equivalent to the squared scaling coefficient.

Generalized Expression of Variance with Shift

$$\mathrm{Var}[D] = \mathrm{Var}[D+a]$$

Generalized Expression of Variance with Scale

$$\mathrm{Var}[\alpha D] = \alpha^2 \mathrm{Var}[D].$$

Example: [Co]Variance of a High Dimensional Dataset.

$D = \{x_1, \ldots, x_n\}, \; x_i \in \mathbb{R}^p$

What happens to the variance if

the dataset was transformed by a linear transformation, e.g, $Ax+b$, for every datapoint.

$$Var[AD+b] = A\,Var[D]\,A^T.$$

(*) This equation represents the covariance matrix.

Since the shifting of datapoints doesn't affect the variance, only the the matrix scaling does. Therefor only have to compute the squares of the scaling elements (Proof @ bottom pg.)

Summary of Linear Transformations Effects on the Mean and Variance

	Mean (μ)	Var (σ^2)
Shifting	$E[D]+a$	$Var[\alpha D] = Var[D]$
Scaling	$E[\alpha D] = \alpha E[D]$	$Var[\alpha D] = \alpha^2 Var[D]$
Both	$E[\alpha D+a] = \alpha E[D]+a$	$\underline{Var[AD+b] = A\,Var[D]\,A^T}$

Pretty freaking cool!

$\boxed{\text{Proof of } Var(Ax) = A(Var(x))A^T}$

$$
\begin{aligned}
Var(AX) &= E[(A(x-\mu))(A(x-\mu))^T] \\
&= E[(A(X-\mu)(X-\mu)^T A^T] \\
&= A\,E[(x-\mu)(x-\mu)^T]\,A^T \\
&= A(Var X)A^T
\end{aligned}
$$

← This can less rigorously be used to prove $Var(Ax+b)$ when $b=\mu$.

• is this just the dot product of each row with every other row to produce the element of $C_{\text{Row, other row}}$.