The quantity

$$\frac{x^\top Ax}{x^\top x}$$

is known as the *Rayleigh ratio* or *Rayleigh–Ritz ratio* (see Section 17.6) and Proposition 23.10 is often known as part of the *Rayleigh–Ritz theorem*.

Proposition 23.10 also holds if A is a Hermitian matrix and if we replace $x^\top Ax$ by x^*Ax and $x^\top x$ by x^*x. The proof is unchanged, since a Hermitian matrix has real eigenvalues and is diagonalized with respect to an orthonormal basis of eigenvectors (with respect to the Hermitian inner product).

We then have the following fundamental result showing how *the SVD of X yields the PCs*:

Theorem 23.11. *(SVD yields PCA) Let X be an $n \times d$ matrix of data points $X_1, \ldots, X_n$, and let μ be the centroid of the X_i's. If $X - \mu = VDU^\top$ is an SVD decomposition of $X - \mu$ and if the main diagonal of D consists of the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d$, then the centered points $Y_1, \ldots, Y_d$, where*

$$Y_k = (X - \mu)u_k = kth\ column\ of\ VD$$

and u_k is the kth column of U, are d principal components of X. Furthermore,

$$\mathrm{var}(Y_k) = \frac{\sigma_k^2}{n-1}$$

and $\mathrm{cov}(Y_h, Y_k) = 0$, whenever $h \neq k$ and $1 \leq k, h \leq d$.

Proof. Recall that for any unit vector v, the centered projection of the points $X_1, \ldots, X_n$ onto the line of direction v is $Y = (X - \mu)v$ and that the variance of Y is given by

$$\mathrm{var}(Y) = v^\top \frac{1}{n-1}(X-\mu)^\top (X-\mu)v.$$

Since $X - \mu = VDU^\top$, we get

$$\begin{aligned}
\mathrm{var}(Y) &= v^\top \frac{1}{(n-1)}(X-\mu)^\top (X-\mu)v \\
&= v^\top \frac{1}{(n-1)} UDV^\top VDU^\top v \\
&= v^\top U \frac{1}{(n-1)} D^2 U^\top v.
\end{aligned}$$

Similarly, if $Y = (X - \mu)v$ and $Z = (X - \mu)w$, then the covariance of Y and Z is given by

$$\mathrm{cov}(Y, Z) = v^\top U \frac{1}{(n-1)} D^2 U^\top w.$$