Let us investigate the effect of a change of variables in the $N\times N$ real & symmetric "***covariance matrix***" *aka* the "***variance matrix***" *aka* the "***error matrix***" $\underline{\hat{V}}_{\underline{x}} \equiv E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T] \equiv E\left[\left(\underline{x}-\underline{\hat{x}}\right)\left(\underline{x}-\underline{\hat{x}}\right)^T\right]$:

$$\underline{\hat{V}}_{\underline{x}} \equiv E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T] \equiv E\left[\left(\underline{x}-E[\underline{x}]\right)\left(\underline{x}-E[\underline{x}]\right)^T\right] \equiv E\left[\left(\underline{x}-\underline{\hat{x}}\right)\left(\underline{x}-\underline{\hat{x}}\right)^T\right] = \begin{pmatrix} \sigma_{x_1}^2 & \mathrm{cov}(x_1,x_2) & \ldots & \mathrm{cov}(x_1,x_N) \\ \mathrm{cov}(x_2,x_1) & \sigma_{x_2}^2 & \ldots & \mathrm{cov}(x_2,x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(x_N,x_1) & \mathrm{cov}(x_N,x_2) & \ldots & \sigma_{x_N}^2 \end{pmatrix}$$

Define:

$$\begin{aligned} u_1 &\equiv u_1\left(x_1,x_{2,}\ldots,x_N\right) \\ u_2 &\equiv u_2\left(x_1,x_{2,}\ldots,x_N\right) \\ &\vdots \\ u_k &\equiv u_k\left(x_1,x_{2,}\ldots,x_N\right) \end{aligned}$$

(n.b. for now, we do not ***require*** that $k\equiv N$)

We can generically define the ***collection*** of u_i vectors as a $k\times 1$ matrix:

$$\underline{u} \equiv \underline{u}\left(\underline{x}\right) = \underbrace{\begin{pmatrix} u_1\left(x_1,x_2,\ldots,x_N\right) \\ u_2\left(x_1,x_2,\ldots,x_N\right) \\ \vdots \\ u_k\left(x_1,x_2,\ldots,x_N\right) \end{pmatrix}}_{\substack{k\times 1 \\ \text{matrix}}}$$

Let us assume that we can expand each $u_i\left(x_1,x_{2,}\ldots,x_N\right)$ vector in a Taylor series about $\underline{\hat{x}}$ and also assume that 2^{nd} order (and all higher order) terms can (safely) be neglected. Then:

$$u_i\left(x_1,x_{2,}\ldots,x_N\right) \cong u_i\left(\hat{x}_1,\hat{x}_{2,}\ldots,\hat{x}_N\right) + \sum_{j=1}^{N} \left.\frac{\partial u_i\left(x_1,x_{2,}\ldots,x_N\right)}{\partial x_j}\right|_{x_j=\hat{x}_j} \left(x_j-\hat{x}_j\right)$$

which we can also symbolically abbreviate as: $\underline{u}\left(\underline{x}\right) \cong \underline{u}\left(\underline{\hat{x}}\right) + \frac{\partial \underline{u}}{\partial \underline{\hat{x}}}\left(\underline{x}-\underline{\hat{x}}\right) = \underline{u}\left(\underline{\hat{x}}\right) + \frac{\partial \underline{u}}{\partial \underline{\hat{x}}}\underline{R}_x$

Explicitly writing this out:

$$\underbrace{\begin{pmatrix} u_1\left(x_1,x_2,\ldots,x_N\right) \\ u_2\left(x_1,x_2,\ldots,x_N\right) \\ \vdots \\ u_k\left(x_1,x_2,\ldots,x_N\right) \end{pmatrix}}_{\substack{k\times 1 \\ \text{matrix}}} = \underbrace{\begin{pmatrix} u_1\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right) \\ u_2\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right) \\ \vdots \\ u_k\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right) \end{pmatrix}}_{\substack{k\times 1 \\ \text{matrix}}} + \underbrace{\underbrace{\begin{pmatrix} \left.\frac{\partial u_1\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right)}{\partial x_1}\right|_{\underline{x}=\underline{\hat{x}}} & \cdots & \left.\frac{\partial u_1\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right)}{\partial x_N}\right|_{\underline{x}=\underline{\hat{x}}} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial u_k\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right)}{\partial x_1}\right|_{\underline{x}=\underline{\hat{x}}} & \cdots & \left.\frac{\partial u_k\left(\hat{x}_1,\hat{x}_2,\ldots,\hat{x}_N\right)}{\partial x_N}\right|_{\underline{x}=\underline{\hat{x}}} \end{pmatrix}}_{\substack{k\times N \\ \text{matrix}}} \underbrace{\begin{pmatrix} x_1-\hat{x}_1 \\ \vdots \\ x_N-\hat{x}_N \end{pmatrix}}_{\substack{N\times 1 \\ \text{matrix}}}}_{\substack{k\times 1 \\ \text{matrix}}}$$

The expectation value of $\underline{u} = \underline{u}(\underline{x})$ is also $E[\underline{u}] = E[\underline{u}(\underline{x})] = \underline{u}(E[\underline{x}]) = \underline{u}(\hat{\underline{x}}) \equiv \hat{\underline{u}}$, as we expect.
We can then determine the $k \times k$ ***covariance*** matrix $\hat{\underline{V}}_{\underline{u}}$ associated with the set of variables $\underline{u} \equiv \underline{u}(\underline{x})$, defined analogously to those we defined for $\hat{\underline{V}}_{\underline{x}}$:

$$\hat{\underline{V}}_{\underline{u}} \equiv E[\underline{R}_{\underline{u}}\underline{R}_{\underline{u}}^T] \equiv E\left[(\underline{u}-E[\underline{u}])(\underline{u}-E[\underline{u}])^T\right] \equiv E\left[(\underline{u}-\hat{\underline{u}})(\underline{u}-\hat{\underline{u}})^T\right] = \begin{pmatrix} \sigma_{u_1}^2 & \mathrm{cov}(u_1,u_2) & \ldots & \mathrm{cov}(u_1,u_k) \\ \mathrm{cov}(u_2,u_1) & \sigma_{u_2}^2 & \ldots & \mathrm{cov}(u_2,u_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(u_k,u_1) & \mathrm{cov}(u_k,u_2) & \ldots & \sigma_{u_k}^2 \end{pmatrix}$$

where the $k \times 1$ $\underline{u}$***-residual*** vector $\underline{R}_{\underline{u}}$ is defined analogously to that for $\underline{R}_{\underline{x}} \equiv \underline{x} - \hat{\underline{x}}$, i.e. $\underline{R}_{\underline{u}} \equiv \underline{u} - \hat{\underline{u}}$.

<u>The details</u>:

$$\underbrace{\underline{R}_{\underline{u}}}_{k\times1} \equiv \underbrace{\underline{u}(\underline{x})}_{k\times1} - \underbrace{E[\underline{u}(\underline{x})]}_{k\times1} = \underbrace{\underline{u}(\underline{x})}_{k\times1} - \underbrace{\underline{u}(E[\underline{x}])}_{k\times1} = \underbrace{\underline{u}(\underline{x})}_{k\times1} - \underbrace{\underline{u}(\hat{\underline{x}})}_{k\times1} = \underbrace{\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)}_{k\times N}\underbrace{(\underline{x}-\hat{\underline{x}})}_{N\times1} = \underbrace{\hat{\underline{D}}_{u/x}}_{k\times N}\underbrace{\underline{R}_x}_{N\times1}$$

where we have also defined the $k \times N$ ***derivative*** matrix $\hat{\underline{D}}_{u/x}$ and its $N \times k$ transpose $\hat{\underline{D}}_{u/x}^T$ as:

$$\hat{\underline{D}}_{u/x} \equiv \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right) = \underbrace{\begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{x}} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{x}} & \cdots & \left.\frac{\partial u_1}{\partial x_N}\right|_{\hat{x}} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{x}} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{x}} & \cdots & \left.\frac{\partial u_2}{\partial x_N}\right|_{\hat{x}} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial u_k}{\partial x_1}\right|_{\hat{x}} & \left.\frac{\partial u_k}{\partial x_2}\right|_{\hat{x}} & \cdots & \left.\frac{\partial u_k}{\partial x_N}\right|_{\hat{x}} \end{pmatrix}}_{k\times N} \quad \text{and} \quad \hat{\underline{D}}_{u/x}^T \equiv \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)^T = \underbrace{\begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{x}} & \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{x}} & \cdots & \left.\frac{\partial u_k}{\partial x_1}\right|_{\hat{x}} \\ \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{x}} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{x}} & \cdots & \left.\frac{\partial u_k}{\partial x_2}\right|_{\hat{x}} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial u_1}{\partial x_N}\right|_{\hat{x}} & \left.\frac{\partial u_2}{\partial x_N}\right|_{\hat{x}} & \cdots & \left.\frac{\partial u_k}{\partial x_N}\right|_{\hat{x}} \end{pmatrix}}_{k\times N}$$

n.b. The $k \times N$ ***derivative*** matrix $\hat{\underline{D}}_{u/x}$ and its $N \times k$ transpose $\hat{\underline{D}}_{u/x}^T$ are ***<u>not</u> symmetric*** matrices!
$\Rightarrow$ For $k \times k$ matrices, in general $\hat{\underline{D}}_{u/x}^T \neq \hat{\underline{D}}_{u/x}$.

Then:

$$\underbrace{\underline{R}_{\underline{u}}}_{k\times1}\underbrace{\underline{R}_{\underline{u}}^T}_{1\times k} \equiv \underbrace{\{\underline{u}(\underline{x}) - E[\underline{u}(\underline{x})]\}}_{k\times1}\underbrace{\{\underline{u}(\underline{x}) - E[\underline{u}(\underline{x})]\}^T}_{1\times k} = \underbrace{\{\underline{u}(\underline{x}) - \underline{u}(\hat{\underline{x}})\}}_{k\times1}\underbrace{\{\underline{u}(\underline{x}) - \underline{u}(\hat{\underline{x}})\}^T}_{1\times k}$$

$$= \underbrace{\left\{\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)(\underline{x}-\hat{\underline{x}})\right\}}_{k\times1}\underbrace{\left\{\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)(\underline{x}-\hat{\underline{x}})\right\}^T}_{1\times k} = \underbrace{\underbrace{\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)}_{k\times N}\underbrace{(\underline{x}-\hat{\underline{x}})}_{N\times1}}_{k\times1}\underbrace{\underbrace{(\underline{x}-\hat{\underline{x}})^T}_{1\times N}\underbrace{\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)^T}_{N\times k}}_{1\times k}$$

$$= \hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\left(\hat{\underline{D}}_{\underline{u}}\underline{R}_{\underline{x}}\right)^T = \hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T\hat{\underline{D}}_{u/x}^T = \text{a } k \times k \text{ matrix.}$$

where: $\underline{R}_{\underline{u}} \equiv \hat{\underline{D}}_{u/x}\underline{R}_x$ and: $\underline{R}_{\underline{u}}^T \equiv \left(\hat{\underline{D}}_{u/x}\underline{R}_x\right)^T = \underline{R}_{\underline{x}}^T\hat{\underline{D}}_{u/x}^T$.

n.b. Recall from linear algebra that the *transpose* of the *product* of two matrices *A* and *B* is $(AB)^T = B^T A^T$.

The $k\times k$ ***covariance*** matrix $\hat{\underline{V}}_u$ associated with the $\underline{u}\equiv\underline{u}(\underline{x})$ variables can thus be written as:

$$\begin{aligned}
\hat{\underline{V}}_{\underline{u}} &\equiv E[\underline{R}_{\underline{u}}\underline{R}_{\underline{u}}^T] = E[\hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\{\hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\}^T] = E[\hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T\hat{\underline{D}}_{u/x}^T] = \hat{\underline{D}}_{u/x}E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T]\hat{\underline{D}}_{u/x}^T \\
&\equiv E\left[\left(\underline{u}-E[\underline{u}]\right)\left(\underline{u}-E[\underline{u}]\right)^T\right] \qquad \equiv E\left[\left(\underline{u}-\hat{\underline{u}}\right)\left(\underline{u}-\hat{\underline{u}}\right)^T\right] \\
&= E\left[\underbrace{\left(\frac{\partial\underline{u}}{\partial\hat{\underline{x}}}\right)}_{k\times N}\underbrace{\left(\underline{x}-\hat{\underline{x}}\right)}_{N\times 1}\underbrace{\left(\underline{x}-\hat{\underline{x}}\right)^T}_{1\times N}\underbrace{\left(\frac{\partial\underline{u}}{\partial\hat{\underline{x}}}\right)^T}_{N\times k}\right] = \underbrace{\left(\frac{\partial\underline{u}}{\partial\hat{\underline{x}}}\right)}_{k\times N}\underbrace{E\left[\underbrace{\left(\underline{x}-\hat{\underline{x}}\right)}_{N\times 1}\underbrace{\left(\underline{x}-\hat{\underline{x}}\right)^T}_{1\times N}\right]}_{N\times N}\underbrace{\left(\frac{\partial\underline{u}}{\partial\hat{\underline{x}}}\right)^T}_{N\times k} \\
&= \begin{pmatrix} \sigma_{u_1}^2 & \mathrm{cov}(u_1,u_2) & \ldots & \mathrm{cov}(u_1,u_k) \\ \mathrm{cov}(u_2,u_1) & \sigma_{u_2}^2 & \ldots & \mathrm{cov}(u_2,u_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(u_k,u_1) & \mathrm{cov}(u_k,u_2) & \ldots & \sigma_{u_k}^2 \end{pmatrix} = \text{ a } k\times k \text{ matrix.}
\end{aligned}$$

Note that in the last step of the 1st and 3rd rows above, we took advantage of the fact that the $k\times N$ ***derivative*** matrix $\hat{\underline{D}}_{u/x}\equiv\left(\frac{\partial\underline{u}}{\partial\hat{\underline{x}}}\right)=\left(\frac{\partial\underline{u}}{\partial\underline{x}}\right)_{x=\hat{x}}$ and its $N\times k$ transpose $\hat{\underline{D}}_{u/x}^T\equiv\left(\frac{\partial\underline{u}}{\partial\hat{\underline{x}}}\right)^T=\left(\frac{\partial\underline{u}}{\partial\underline{x}}\right)_{x=\hat{x}}^T$ are just sets of ***<u>numbers</u>*** (constants!), i.e. they are no longer ***functions*** of the random variables x_i .

Now, since the $N\times N$ ***covariance matrix*** $\hat{\underline{V}}_{\underline{x}}\equiv E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T]\equiv E\left[\left(\underline{x}-\hat{\underline{x}}\right)\left(\underline{x}-\hat{\underline{x}}\right)^T\right]$ is:

$$\begin{aligned}
\hat{\underline{V}}_{\underline{x}} &\equiv E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T]\equiv E\left[\left(\underline{x}-E[\underline{x}]\right)\left(\underline{x}-E[\underline{x}]\right)^T\right]\equiv E\left[\left(\underline{x}-\hat{\underline{x}}\right)\left(\underline{x}-\hat{\underline{x}}\right)^T\right] \\
&= \begin{pmatrix} \sigma_{x_1}^2 & \mathrm{cov}(x_1,x_2) & \ldots & \mathrm{cov}(x_1,x_N) \\ \mathrm{cov}(x_2,x_1) & \sigma_{x_2}^2 & \ldots & \mathrm{cov}(x_2,x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(x_N,x_1) & \mathrm{cov}(x_N,x_2) & \ldots & \sigma_{x_N}^2 \end{pmatrix}
\end{aligned}$$

The $k\times k$ ***covariance matrix*** $\hat{\underline{V}}_{\underline{u}}$ can then be compactly written as:

$$\hat{\underline{V}}_{\underline{u}}\equiv E[\underline{R}_{\underline{u}}\underline{R}_{\underline{u}}^T]=E[\hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\{\hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\}^T]=E[\hat{\underline{D}}_{u/x}\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T\hat{\underline{D}}_{u/x}^T]=\hat{\underline{D}}_{u/x}\underbrace{E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T]}_{\equiv\hat{\underline{V}}_{\underline{x}}}\hat{\underline{D}}_{u/x}^T=\hat{\underline{D}}_{u/x}\hat{\underline{V}}_{\underline{x}}\hat{\underline{D}}_{u/x}^T$$

<u>Important note</u>:

In various probability & statistics textbooks, the covariance matrix equation is instead written as $\hat{\underline{V}}_{\underline{u}}=\hat{\underline{D}}_{u/x}^{\prime T}\hat{\underline{V}}_{\underline{x}}\hat{\underline{D}}'_{u/x}$. A detailed comparison of $\hat{\underline{V}}_{\underline{u}}=\hat{\underline{D}}_{u/x}^{\prime T}\hat{\underline{V}}_{\underline{x}}\hat{\underline{D}}'_{u/x}$ with the above derivation of $\hat{\underline{V}}_{\underline{u}}=\hat{\underline{D}}_{u/x}\hat{\underline{V}}_{\underline{x}}\hat{\underline{D}}_{u/x}^T$ shows that the two seemingly different/contradictory relations ***<u>are</u>*** indeed equivalent, because $\hat{\underline{D}}_{u/x}=\hat{\underline{D}}_{u/x}^{\prime T}$ and $\hat{\underline{D}}_{u/x}^T=\hat{\underline{D}}'_{u/x}$! The origin of the two differing conventions can be traced back to the respective ***<u>definitions</u>*** of the ***residual*** matrices as a ***column*** vector ($N\times 1$ matrix) $\underline{R}_{\underline{x}}\equiv\left(\underline{x}-\hat{\underline{x}}\right)$ *vs*. a ***row*** vector ($1\times N$ matrix) $\underline{R}'_{\underline{x}}\equiv\left(\underline{x}-\hat{\underline{x}}\right)$ in the formation of the ***outer product*** for the $N\times N$ covariance matrix $\hat{\underline{V}}_{\underline{x}}\equiv E[\underline{R}_{\underline{x}}\underline{R}_{\underline{x}}^T]=E[\underline{R}_{\underline{x}}^{\prime T}\underline{R}'_{\underline{x}}]$, i.e. $\underline{R}_{\underline{x}}=\underline{R}_{\underline{x}}^{\prime T}$ and $\underline{R}_{\underline{x}}^T=\underline{R}'_{\underline{x}}$.

Explicitly writing out this relation, long-hand, it is:

$$\underline{\hat{V}}_{\underline{u}} = \underbrace{\begin{pmatrix} \sigma_{u_1}^2 & \operatorname{cov}(u_1,u_2) & \dots & \operatorname{cov}(u_1,u_k) \\ \operatorname{cov}(u_2,u_1) & \sigma_{u_2}^2 & \dots & \operatorname{cov}(u_2,u_k) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}(u_k,u_1) & \operatorname{cov}(u_k,u_2) & \dots & \sigma_{u_k}^2 \end{pmatrix}}_{k\times k}$$

$$= \underbrace{\begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_1}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_2}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial u_k}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_k}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_k}{\partial x_N}\right|_{\hat{\underline{x}}} \end{pmatrix}}_{k\times N} \underbrace{\begin{pmatrix} \sigma_{x_1}^2 & \operatorname{cov}(x_1,x_2) & \dots & \operatorname{cov}(x_1,x_N) \\ \operatorname{cov}(x_2,x_1) & \sigma_{x_2}^2 & \dots & \operatorname{cov}(x_2,x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}(x_N,x_1) & \operatorname{cov}(x_N,x_2) & \dots & \sigma_{x_N}^2 \end{pmatrix}}_{N\times N} \underbrace{\begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_1}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_2}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial u_k}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_k}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_k}{\partial x_N}\right|_{\hat{\underline{x}}} \end{pmatrix}^T}_{N\times k}$$

For the common case of N ***independent*** variables $x_1, x_2, \cdots, x_N$ ***.and.*** ***N functions*** $u_1, u_2, \cdots, u_N$ all of the ***covariances*** vanish, thus things simplify, and we then have, in explicit, long-hand notation:

$$\underline{\hat{V}}_{\underline{x}} = \underbrace{\begin{pmatrix} \sigma_{x_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{x_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{x_N}^2 \end{pmatrix}}_{N\times N}$$

So:

$$\underline{\hat{V}}_{\underline{u}} = \underbrace{\begin{pmatrix} \sigma_{u_1}^2 & \operatorname{cov}(u_1,u_2) & \dots & \operatorname{cov}(u_1,u_N) \\ \operatorname{cov}(u_2,u_1) & \sigma_{u_2}^2 & \dots & \operatorname{cov}(u_2,u_N) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cov}(u_N,u_1) & \operatorname{cov}(u_N,u_2) & \dots & \sigma_{u_N}^2 \end{pmatrix}}_{N\times N}$$

$$= \underbrace{\begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_1}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_2}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial u_k}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_k}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_k}{\partial x_N}\right|_{\hat{\underline{x}}} \end{pmatrix}}_{N\times N} \underbrace{\begin{pmatrix} \sigma_{x_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{x_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{x_N}^2 \end{pmatrix}}_{N\times N} \underbrace{\begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_1}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_2}{\partial x_N}\right|_{\hat{\underline{x}}} \\ \vdots & \vdots & \ddots & \vdots \\ \left.\frac{\partial u_k}{\partial x_1}\right|_{\hat{\underline{x}}} & \left.\frac{\partial u_k}{\partial x_2}\right|_{\hat{\underline{x}}} & \cdots & \left.\frac{\partial u_k}{\partial x_N}\right|_{\hat{\underline{x}}} \end{pmatrix}^T}_{N\times N}$$

For the case of N ***independent*** random variables $x_1, x_2, \cdots, x_N$ ***.and. N functions*** $u_1, u_2, \cdots, u_N$ the $N \times N$ ***covariance*** matrix $\hat{\underline{V}}_u$ has diagonal elements of the form:

$$\left(\hat{\underline{V}}_{\underline{u}}\right)_{kk} = \text{cov}\left(u_k, u_k\right) = \text{var}\left(u_k\right) = \sigma_{u_k}^2 = \sum_{i=1}^{N}\left(\left.\frac{\partial u_k\left(x_i\right)}{\partial x_i}\right|_{x_i=\hat{x}_i}\right)^2 \sigma_{x_i}^2 \quad \text{for} \quad k = 1, 2, \ldots, N$$

which agrees with the corresponding Taylor series derivation on p. 5 of P598AEM Lect. Notes 5.

In addition, the $N \times N$ ***covariance*** matrix $\hat{\underline{V}}_u$ ***also*** has non-zero ***off***-diagonal elements of the form:

$$\left(\hat{\underline{V}}_{\underline{u}}\right)_{k\ell} = \text{cov}\left(u_k, u_l\right) = \sum_{i=1}^{N}\left(\left.\frac{\partial u_k\left(x_i\right)}{\partial x_i}\right|_{x_i=\hat{x}_i}\right)\sigma_{x_i}^2\left(\left.\frac{\partial u_\ell\left(x_i\right)}{\partial x_i}\right|_{x_i=\hat{x}_i}\right) \quad \text{for} \quad k \neq \ell = 1, 2, \ldots, N$$

which need ***not*** be zero! The transformation from $x \to u$ has ***induced*** correlations between the ***new*** random variables u_i even though the ***original*** x_i variables were ***independent***!!!

A simple example of $\hat{\underline{V}}_{\underline{u}} = \hat{\underline{D}}_{u/x}\hat{\underline{V}}_{\underline{x}}\hat{\underline{D}}_{u/x}^T = \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)\hat{\underline{V}}_{\underline{x}}\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)^T$:

Let x_1 and x_2 be ***independent*** random variables. Let $u_1 \equiv \frac{1}{\sqrt{2}}\left(x_1 - x_2\right)$ and $u_2 \equiv \frac{1}{\sqrt{2}}\left(x_1 + x_2\right)$.

Assume that we are given the $\underline{x}$-basis ***covariance*** matrix $\hat{\underline{V}}_{\underline{x}} = \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix}$

Then: $$\hat{\underline{D}}_{u/x} = \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right) = \begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{x}_1} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{x}_2} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{x}_1} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{x}_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

And: $$\hat{\underline{D}}_{u/x}^T = \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)^T = \begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{x}_1} & \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{x}_2} \\ \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{x}_1} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{x}_2} \end{pmatrix}^T = \begin{pmatrix} \left.\frac{\partial u_1}{\partial x_1}\right|_{\hat{x}_1} & \left.\frac{\partial u_2}{\partial x_1}\right|_{\hat{x}_1} \\ \left.\frac{\partial u_1}{\partial x_2}\right|_{\hat{x}_2} & \left.\frac{\partial u_2}{\partial x_2}\right|_{\hat{x}_2} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right) \neq \hat{\underline{D}}_u \ \text{(here)}$$

So:

$$\hat{\underline{V}}_{\underline{u}} = \hat{\underline{D}}_{u/x}\hat{\underline{V}}_{\underline{x}}\hat{\underline{D}}_{u/x}^T = \left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)\hat{\underline{V}}_{\underline{x}}\left(\frac{\partial \underline{u}}{\partial \hat{\underline{x}}}\right)^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} \sigma_{x_1}^2 + \sigma_{x_2}^2 & \sigma_{x_1}^2 - \sigma_{x_2}^2 \\ \sigma_{x_1}^2 - \sigma_{x_2}^2 & \sigma_{x_1}^2 + \sigma_{x_2}^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right) & \frac{1}{2}\left(\sigma_{x_1}^2 - \sigma_{x_2}^2\right) \\ \frac{1}{2}\left(\sigma_{x_1}^2 - \sigma_{x_2}^2\right) & \frac{1}{2}\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right) \end{pmatrix} = \begin{pmatrix} \sigma_{u_1}^2 & \text{cov}\left(u_1, u_2\right) \\ \text{cov}\left(u_2, u_1\right) & \sigma_{u_2}^2 \end{pmatrix}$$

Thus in ***general*** u_1 and u_2 are ***not independent*** {b/c $\text{cov}\left(u_1, u_2\right) = \frac{1}{2}\left(\sigma_{x_1}^2 - \sigma_{x_2}^2\right) \neq 0$ } ***unless*** $\sigma_{x_1} \equiv \sigma_{x_2}$.

A Numerical Example:

Suppose the ***1-sigma uncertainty*** on x_1 is $\sigma_{x_1} = 5$ and the ***1-sigma uncertainty*** on x_2 is $\sigma_{x_2} = 3$.

The 2×2 ***covariance*** matrix associated with the transformation $u_1 \equiv \frac{1}{\sqrt{2}}(x_1 - x_2)$, $u_2 \equiv \frac{1}{\sqrt{2}}(x_1 + x_2)$ is:

$$\underline{\hat{V}}_u = \begin{pmatrix} \sigma_{u_1}^2 & \mathrm{cov}(u_1,u_2) \\ \mathrm{cov}(u_2,u_1) & \sigma_{u_2}^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\sigma_{x_1}^2 + \sigma_{x_2}^2) & \frac{1}{2}(\sigma_{x_1}^2 - \sigma_{x_2}^2) \\ \frac{1}{2}(\sigma_{x_1}^2 - \sigma_{x_2}^2) & \frac{1}{2}(\sigma_{x_1}^2 + \sigma_{x_2}^2) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(25+9) & \frac{1}{2}(25-9) \\ \frac{1}{2}(25-9) & \frac{1}{2}(25+9) \end{pmatrix} = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$$

Thus the ***1-sigma uncertainty*** on $u_1 = \frac{1}{\sqrt{2}}(x_1 + x_2)$ is: $\sigma_{u_1} = \sqrt{17} = 4.1231$

and the ***1-sigma uncertainty*** on $u_2 = \frac{1}{\sqrt{2}}(x_1 - x_2)$ is: $\sigma_{u_2} = \sqrt{17} = 4.1231 = \sigma_{u_1}$

However, even though (here) $\sigma_{u_1} \equiv \sigma_{u_2} = \sqrt{17} = 4.1231$, the random variables u_1 and u_2 are ***not*** independent – they are (positively) correlated with each other, because:

$$\mathrm{cov}(u_1,u_2) = \mathrm{cov}(u_2,u_1) = \tfrac{1}{2}(\sigma_{x_1}^2 - \sigma_{x_2}^2) = 8 > 0$$

or, equivalently:

$$\rho(u_1,u_2) = \frac{\mathrm{cov}(u_1,u_2)}{\sigma_{u_1}\sigma_{u_2}} = \frac{8}{17} = 0.4706\,.$$

Undoing the Error Propagation:

Inverting this transformation, we have: $x_1 \equiv \frac{1}{\sqrt{2}}(u_1 + u_2)$ and $x_2 \equiv \frac{1}{\sqrt{2}}(-u_1 + u_2)$.

Let us (temporarily) pretend to ***ignore*** correlations, i.e. treat u_1 and u_2 as if they ***were*** independent.

We ***would*** then obtain:

$$\sigma_{x_1}^2 = \left(\frac{\partial x_1}{\partial \hat{u}_1}\right)^2 \sigma_{u_1}^2 + \left(\frac{\partial x_1}{\partial \hat{u}_2}\right)^2 \sigma_{u_2}^2 = \tfrac{1}{2}\cdot 17 + \tfrac{1}{2}\cdot 17 = 17 \quad \Rightarrow \quad \sigma_{x_1} = \sqrt{17} = 4.1231$$

$$\sigma_{x_2}^2 = \left(\frac{\partial x_2}{\partial \hat{u}_1}\right)^2 \sigma_{u_1}^2 + \left(\frac{\partial x_2}{\partial \hat{u}_2}\right)^2 \sigma_{u_2}^2 = \tfrac{1}{2}\cdot 17 + \tfrac{1}{2}\cdot 17 = 17 \quad \Rightarrow \quad \sigma_{x_2} = \sqrt{17} = 4.1231 = \sigma_{x_1}$$

These results are clearly *wrong* !!! (Since we initially stated that $\sigma_{x_1} = 5$ and $\sigma_{x_2} = 3$ **!!!**)

The (correct) ***inverse*** transformation, using matrices is: $\underline{\hat{V}}_x \equiv \underline{D}_{x/u}\underline{\hat{V}}_u\underline{D}_{x/u}^T \equiv \left(\frac{\partial \underline{x}}{\partial \underline{\hat{u}}}\right)\underline{\hat{V}}_u\left(\frac{\partial \underline{x}}{\partial \underline{\hat{u}}}\right)^T$

where:

$$\underline{D}_{x/u} \equiv \left(\frac{\partial \underline{x}}{\partial \underline{\hat{u}}}\right) = \begin{pmatrix} \frac{\partial x_1}{\partial \hat{u}_1} & \frac{\partial x_1}{\partial \hat{u}_2} \\ \frac{\partial x_2}{\partial \hat{u}_1} & \frac{\partial x_2}{\partial \hat{u}_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \Rightarrow \underline{D}_{x/u}^T \equiv \left(\frac{\partial \underline{x}}{\partial \underline{\hat{u}}}\right)^T = \begin{pmatrix} \frac{\partial x_1}{\partial \hat{u}_1} & \frac{\partial x_2}{\partial \hat{u}_1} \\ \frac{\partial x_1}{\partial \hat{u}_2} & \frac{\partial x_2}{\partial \hat{u}_2} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \neq \underline{D}_{x/u}$$

Then:

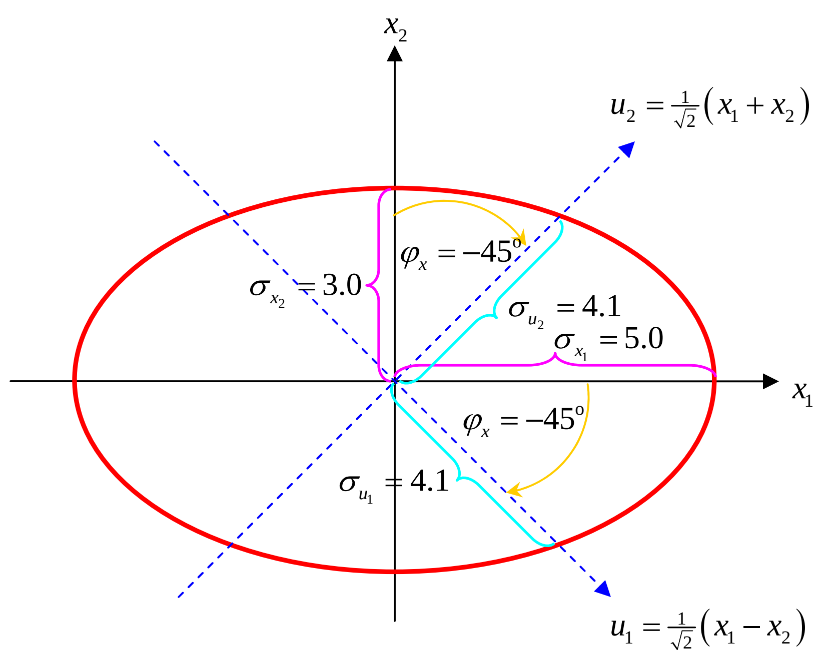
$$\underline{\hat{V}}_{\underline{x}} = \underline{D}_{x/u}\underline{\hat{V}}_{\underline{u}}\underline{D}_{x/u}^{T} = \left(\frac{\partial \underline{x}}{\partial \underline{\hat{u}}}\right)\underline{\hat{V}}_{\underline{u}}\left(\frac{\partial \underline{x}}{\partial \underline{\hat{u}}}\right)^{T} = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

$$= \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 25 & -9 \\ 25 & 9 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 50 & 0 \\ 0 & 18 \end{pmatrix} = \begin{pmatrix} 25 & 0 \\ 0 & 9 \end{pmatrix} = \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix}$$

Thus, we see via the use of the above matrix formalism, that we indeed (correctly) recover the original $\sigma_{x_1} = 5$ and $\sigma_{x_2} = 3$, i.e. the random variables x_1 and x_2 ***are independent***.

The above simple example(s) of a change of (***orthonormal***) variables can be easily understood as a simple change of ***orthonormal basis vectors*** – from the ***independent*** random variables x_1 and x_2 to the ***non-independent*** random variables u_1 and u_2 via the ***orthonormal transformation*** $u_1 \equiv \tfrac{1}{\sqrt{2}}(x_1 - x_2)$, $u_2 \equiv \tfrac{1}{\sqrt{2}}(x_1 + x_2)$ and/or the ***inverse orthonormal transformation*** $x_1 \equiv \tfrac{1}{\sqrt{2}}(u_1 + u_2)$, $x_2 \equiv \tfrac{1}{\sqrt{2}}(-u_1 + u_2)$.
(n.b. both of these orthonormal transformations are from RH → RH coordinate systems…)

The $x \to u$ orthonormal transformation is a consequence of applying a $\varphi_x = -45^{\circ}$ (CW) rotation in the $x_1 - x_2$ plane, as shown by the red ellipse in the figure below:



n.b. we also see from the elliptical symmetry associated with the above figure that $\sigma_{u_1} = \sigma_{u_2} \cong 4.1$ simply because of the ***specific*** choice of the $\varphi = -45^{\circ}$ (CW) rotation in the 2-D $x_1 - x_2$ plane, resulting in the $u_1 - u_2$ basis vectors each having ***equal*** projections onto the $x_1 - x_2$ basis vectors. Had we instead chosen an ***arbitrary*** φ-rotation in the 2-D $x_1 - x_2$ plane, then in general $\sigma_{u_1} \neq \sigma_{u_2}$.

For an ***arbitrary*** φ_x -rotation of $\underline{x}$ -basis vectors:

$$\begin{aligned} u_1 &= x_1\cos\varphi_x + x_2\sin\varphi_x \\ u_2 &= -x_1\sin\varphi_x + x_2\cos\varphi_x \end{aligned} \Rightarrow \underbrace{\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}}_{\equiv \underline{u}} = \begin{pmatrix} x_1\cos\varphi_x + x_2\sin\varphi_x \\ -x_1\sin\varphi_x + x_2\cos\varphi_x \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\varphi_x & \sin\varphi_x \\ -\sin\varphi_x & \cos\varphi_x \end{pmatrix}}_{\equiv \underline{R}_x} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\equiv \underline{x}} \Rightarrow \underline{u} = \underline{R}_x\,\underline{x}$$

The ***inverse*** φ_u -rotation transformation of $\underline{u}$ -basis vectors is (also) given by:

$$\begin{aligned} x_1 &= u_1\cos\varphi_u + u_2\sin\varphi_u \\ x_2 &= -u_1\sin\varphi_u + u_2\cos\varphi_u \end{aligned} \Rightarrow \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\equiv \underline{x}} = \begin{pmatrix} u_1\cos\varphi_u + u_2\sin\varphi_u \\ -u_1\sin\varphi_u + u_2\cos\varphi_u \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\varphi_u & \sin\varphi_u \\ -\sin\varphi_u & \cos\varphi_u \end{pmatrix}}_{\equiv \underline{R}_u^{-1}} \underbrace{\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}}_{\equiv \underline{u}} \Rightarrow \underline{x} = \underline{R}_u^{-1}\underline{u}$$

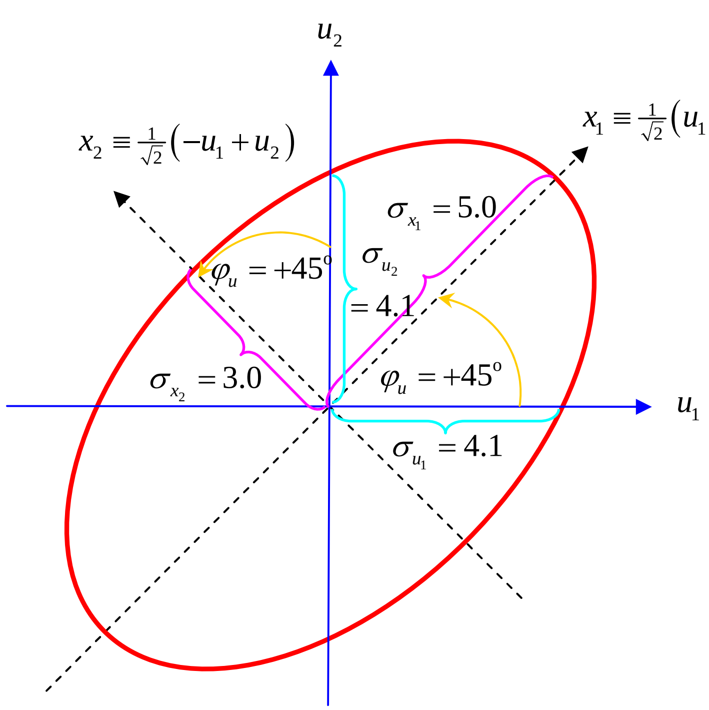
But note that $\varphi_u \equiv -\varphi_x \left(= +45^\circ \text{ here}\right)$, thus we can re-write the ***inverse*** φ_u -rotation transformation as:

$$\begin{aligned} x_1 &= u_1\cos\varphi_x - u_2\sin\varphi_x \\ x_2 &= u_1\sin\varphi_x + u_2\cos\varphi_x \end{aligned} \Rightarrow \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\equiv \underline{x}} = \begin{pmatrix} u_1\cos\varphi_x - u_2\sin\varphi_x \\ u_1\sin\varphi_x + u_2\cos\varphi_x \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\varphi_x & -\sin\varphi_x \\ \sin\varphi_x & \cos\varphi_x \end{pmatrix}}_{\equiv \underline{R}_x^{-1}} \underbrace{\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}}_{\equiv \underline{u}} \Rightarrow \underline{x} = \underline{R}_x^{-1}\underline{u}$$

with:

$$\underline{R}_x\underline{R}_x^{-1} = \begin{pmatrix} \cos\varphi_x & \sin\varphi_x \\ -\sin\varphi_x & \cos\varphi_x \end{pmatrix}\begin{pmatrix} \cos\varphi_x & -\sin\varphi_x \\ \sin\varphi_x & \cos\varphi_x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{\mathbb{1}} = \begin{pmatrix} \cos\varphi_x & -\sin\varphi_x \\ \sin\varphi_x & \cos\varphi_x \end{pmatrix}\begin{pmatrix} \cos\varphi_x & \sin\varphi_x \\ -\sin\varphi_x & \cos\varphi_x \end{pmatrix} = \underline{R}_x^{-1}\underline{R}_x$$

Graphically, the ***inverse*** $u \to x$ orthonormal transformation for a $\varphi_u = +45^\circ$ (CCW) rotation is shown by the red ellipse in the figure below:



The equation for the red ellipse, expressed in terms of the ***independent*** random variable $x_1 - x_2$ ***orthonormal*** basis vectors is given by:

$$\frac{x_1^2}{\sigma_{x_1}^2} + \frac{x_2^2}{\sigma_{x_2}^2} = 1 \quad \Rightarrow \quad \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1/\sigma_{x_1}^2 & 0 \\ 0 & 1/\sigma_{x_2}^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1$$

Note that the matrix $\begin{pmatrix} 1/\sigma_{x_1}^2 & 0 \\ 0 & 1/\sigma_{x_2}^2 \end{pmatrix}$ is the ***inverse*** of the $\underline{x}$-basis ***variance*** matrix $\hat{\underline{V}}_{\underline{x}} = \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix}$:

$$\begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix} \begin{pmatrix} 1/\sigma_{x_1}^2 & 0 \\ 0 & 1/\sigma_{x_2}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{\mathbb{1}} = \begin{pmatrix} 1/\sigma_{x_1}^2 & 0 \\ 0 & 1/\sigma_{x_2}^2 \end{pmatrix} \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix}$$

Which can be written compactly as: $\hat{\underline{V}}_{\underline{x}} \hat{\underline{V}}_{\underline{x}}^{-1} = \underline{\mathbb{1}} = \hat{\underline{V}}_{\underline{x}}^{-1} \hat{\underline{V}}_{\underline{x}}$ where $\underline{\mathbb{1}}$ is the unit matrix, and thus we can also write the above $\underline{x}$-basis ellipse equation compactly as:

$$\frac{x_1^2}{\sigma_{x_1}^2} + \frac{x_2^2}{\sigma_{x_2}^2} = 1 \quad \Rightarrow \quad \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1/\sigma_{x_1}^2 & 0 \\ 0 & 1/\sigma_{x_2}^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1 \quad \Rightarrow \quad \underline{x}^T \hat{\underline{V}}_{\underline{x}}^{-1} \underline{x} = 1$$

The equation for the red ellipse, expressed in terms of the ***non-independent*** random variable $u_1 - u_2$ ***orthonormal*** basis vectors is given by:

$$\frac{u_1^2}{\sigma_{u_1}^2} + \frac{u_2^2}{\sigma_{u_2}^2} - \frac{2\, u_1 u_2 \rho(u_1, u_2)}{\sigma_{u_1}\sigma_{u_2}} = 1 \quad \Rightarrow \quad \begin{pmatrix} u_1 & u_2 \end{pmatrix} \frac{1}{1-\rho^2(u_1,u_2)} \begin{pmatrix} \dfrac{1}{\sigma_{u_1}^2} & -\dfrac{\rho(u_1,u_2)}{\sigma_{u_1}\sigma_{u_2}} \\ -\dfrac{\rho(u_1,u_2)}{\sigma_{u_1}\sigma_{u_2}} & \dfrac{1}{\sigma_{u_2}^2} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 1$$

$$\text{where} \quad \rho(u_1, u_2) \equiv \frac{\mathrm{cov}(u_1, u_2)}{\sigma_{u_1}\sigma_{u_2}}$$

Here, the matrix $\hat{\underline{V}}_{\underline{u}}^{-1} \equiv \dfrac{1}{1-\rho^2(u_1,u_2)} \begin{pmatrix} 1/\sigma_{u_1}^2 & -\rho(u_1,u_2)/\sigma_{u_1}\sigma_{u_2} \\ -\rho(u_1,u_2)/\sigma_{u_1}\sigma_{u_2} & 1/\sigma_{u_2}^2 \end{pmatrix}$ is the ***inverse*** of the $\underline{u}$-basis ***covariance*** matrix $\hat{\underline{V}}_{\underline{u}} = \begin{pmatrix} \sigma_{u_1}^2 & \mathrm{cov}(u_1,u_2) \\ \mathrm{cov}(u_2,u_1) & \sigma_{u_2}^2 \end{pmatrix} = \begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1}\sigma_{u_2}\rho(u_1,u_2) \\ \sigma_{u_2}\sigma_{u_1}\rho(u_2,u_1) & \sigma_{u_2}^2 \end{pmatrix}$:

$$\begin{pmatrix} \sigma_{u_1}^2 & \sigma_{u_1}\sigma_{u_2}\rho(u_1,u_2) \\ \sigma_{u_1}\sigma_{u_2}\rho(u_1,u_2) & \sigma_{u_2}^2 \end{pmatrix} \frac{1}{1-\rho^2(u_1,u_2)} \begin{pmatrix} 1/\sigma_{u_1}^2 & -\rho(u_1,u_2)/\sigma_{u_1}\sigma_{u_2} \\ -\rho(u_1,u_2)/\sigma_{u_1}\sigma_{u_2} & 1/\sigma_{u_2}^2 \end{pmatrix}$$

$$= \frac{1}{1-\rho^2(u_1,u_2)} \begin{pmatrix} 1-\rho^2(u_1,u_2) & \cancel{-\sigma_{u_1}\rho(u_1,u_2)/\sigma_{u_2} + \sigma_{u_1}\rho(u_1,u_2)/\sigma_{u_2}} \\ \cancel{\sigma_{u_2}\rho(u_1,u_2)/\sigma_{u_1} - \sigma_{u_2}\rho(u_1,u_2)/\sigma_{u_1}} & 1-\rho^2(u_1,u_2) \end{pmatrix}$$

$$= \frac{1}{1-\rho^2(u_1,u_2)} \begin{pmatrix} 1-\rho^2(u_1,u_2) & 0 \\ 0 & 1-\rho^2(u_1,u_2) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{\mathbb{1}}$$

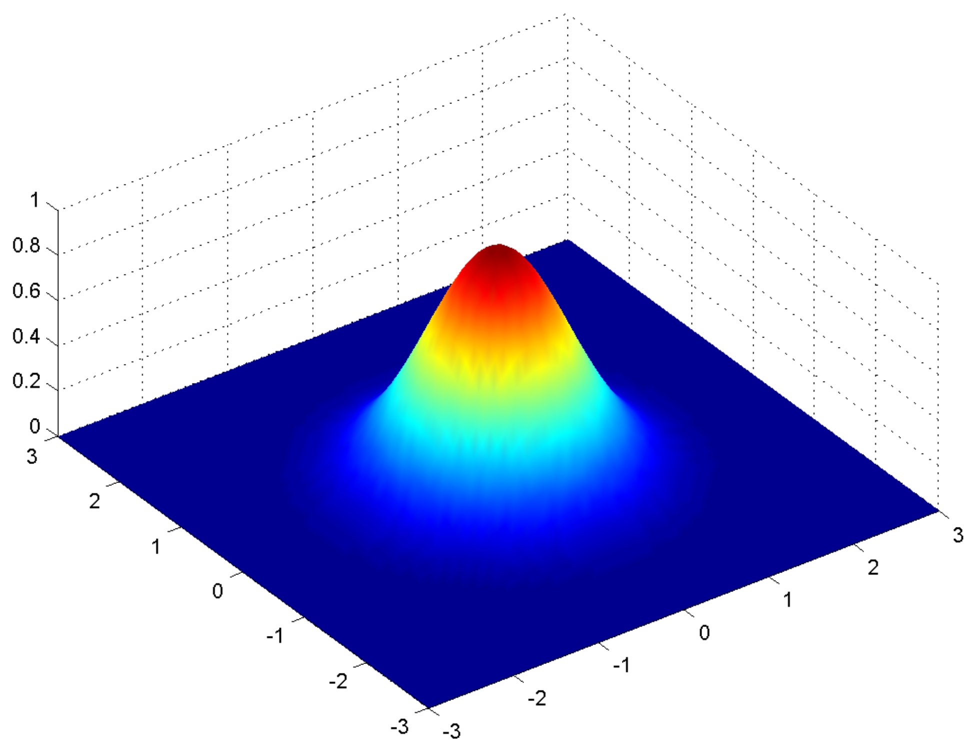
which can be written compactly as: $\underline{\hat{V}}_{\underline{u}}\underline{\hat{V}}_{\underline{u}}^{-1} = \underline{\mathbb{1}} = \underline{\hat{V}}_{\underline{u}}^{-1}\underline{\hat{V}}_{\underline{u}}$ and thus we can also write the above $\underline{u}$-basis ellipse equation compactly as:

$$\frac{u_1^2}{\sigma_{u_1}^2} + \frac{u_2^2}{\sigma_{u_2}^2} - \frac{2\,u_1 u_2 \rho(u_1,u_2)}{\sigma_{u_1}\sigma_{u_2}} = 1 \quad \Rightarrow \quad \underline{u}^T \underline{\hat{V}}_{\underline{u}}^{-1} \underline{u} = 1$$

If the P.D.F. associated with the ***independent*** random variables x_1 and x_2 is the Gaussian/normal distribution, *i.e*.

$$G(x_1,x_2) = G(x_1)\cdot G(x_2) = \frac{1}{\sqrt{2\pi}\,\sigma_{x_1}} e^{-\frac{x_1^2}{2\sigma_{x_1}^2}} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_{x_2}} e^{-\frac{x_2^2}{2\sigma_{x_2}^2}} = \frac{1}{2\pi\,\sigma_{x_1}\sigma_{x_2}} e^{-\left[\frac{x_1^2}{2\sigma_{x_1}^2} + \frac{x_2^2}{2\sigma_{x_2}^2}\right]}$$

The 3-D surface associated with the 2-D Gaussian/normal probability distribution $G(x_1,x_2)$ is shown in the figure below (for the special/limiting case of $\sigma_{x_1} = \sigma_{x_2}$):

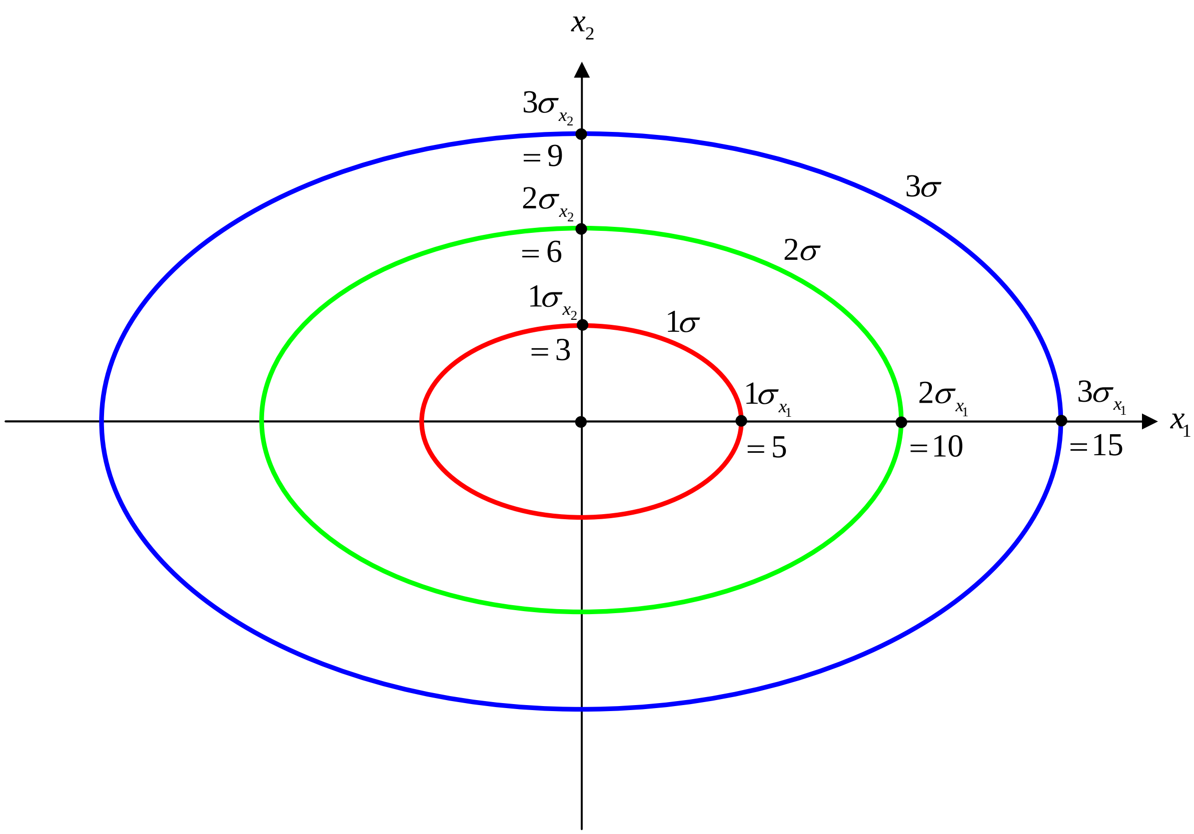


We see that contours of constant/equal probability density are (in general) ellipses in the 2-D $x_1 - x_2$ plane, where the argument of the exponential is equal to a constant, i.e.

$$\left[\frac{x_1^2}{2\sigma_{x_1}^2}+\frac{x_2^2}{2\sigma_{x_2}^2}\right]=\text{constant}$$

At the points $(x_1, x_2) = (\pm\sigma_{x_1}, 0)$ and/or $(x_1, x_2) = (0, \pm\sigma_{x_2})$ on the ellipse curve in the 2-D $x_1 - x_2$ plane, we see that constant = 1/2 (true for any value of (x_1, x_2) lying on this particular ellipse) and thus we see that the equation of an ellipse for which $\left[\frac{x_1^2}{2\sigma_{x_1}^2}+\frac{x_2^2}{2\sigma_{x_2}^2}\right]=\frac{1}{2}$ associated with the above 2-D Gaussian/normal P.D.F. $G(x_1, x_2)$, corresponds to a 1-sigma (= 1 standard deviation) contour (*aka* 1-sigma "equipotential") of constant/equal probability density.

Similarly, e.g. for an arbitrary # of (integer) sigma/standard deviations, i.e. $n_\sigma = 1, 2, 3, 4, 5, \ldots$ this corresponds to contours associated with equation of ellipses for which $\left[\frac{x_1^2}{2\sigma_{x_1}^2}+\frac{x_2^2}{2\sigma_{x_2}^2}\right]=\frac{n_\sigma^2}{2}$ is satisfied, as shown in the figure below:



For the Gaussian/normal P.D.F. associated with the ***dependent*** random variables $u_1 \equiv \frac{1}{\sqrt{2}}\left(x_1 - x_2\right)$ and $u_2 \equiv \frac{1}{\sqrt{2}}\left(x_1 + x_2\right)$:

$$G\left(u_1,u_2\right) = \frac{1}{2\pi\,\sigma_{u_1}\sigma_{u_2}}\frac{1}{\sqrt{1-\rho^2\left(u_1,u_2\right)}}\,e^{-\left[\frac{1}{2}\frac{1}{\left[1-\rho^2\left(u_1,u_2\right)\right]}\left(\frac{u_1^2}{\sigma_{u_1}^2}+\frac{u_2^2}{\sigma_{u_2}^2}-\frac{2\,u_1u_2\rho\left(u_1,u_2\right)}{\sigma_{u_1}\sigma_{u_2}}\right)\right]} \quad \text{where} \quad \rho\left(u_1,u_2\right) \equiv \frac{\operatorname{cov}\left(u_1,u_2\right)}{\sigma_{u_1}\sigma_{u_2}}$$

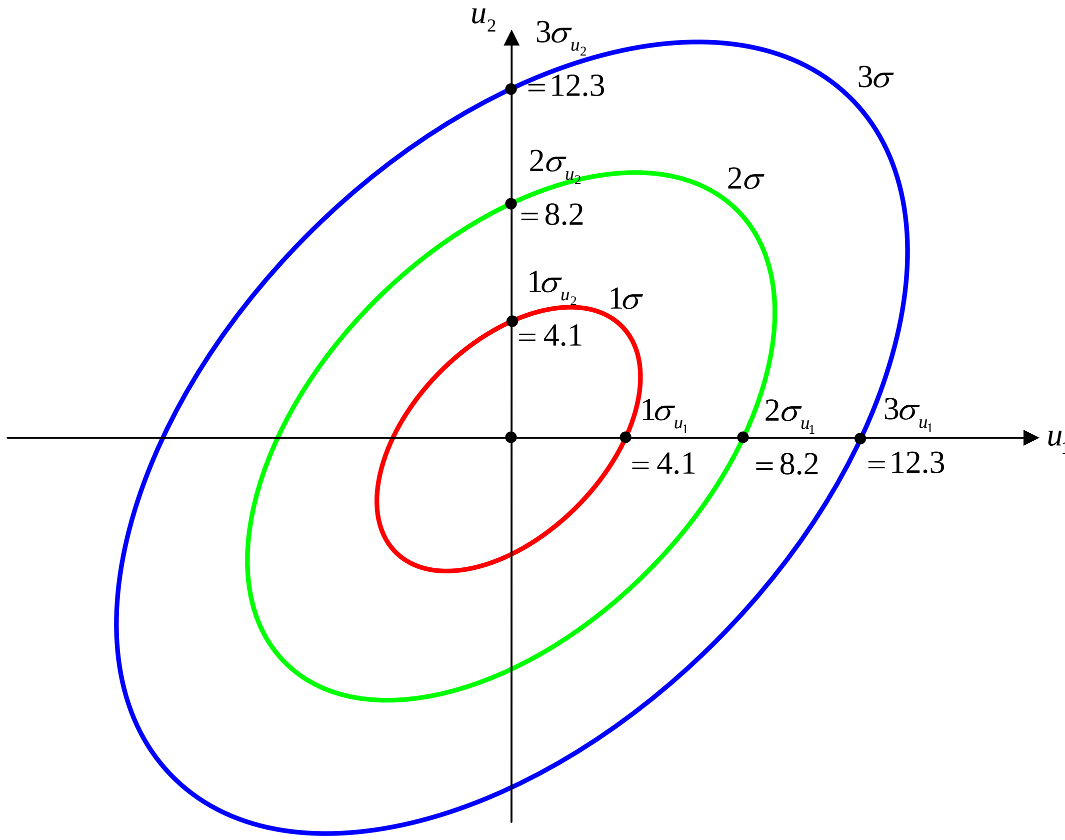
{This is the general form of a 2-D Gaussian distribution in two variables, including correlations}

Then we see that contours of constant/equal probability density are indeed ellipses in the 2-D $u_1 - u_2$ plane, where the argument of the exponential is equal to a constant, i.e.

$$\left[\frac{1}{2}\frac{1}{\left[1-\rho^2\left(u_1,u_2\right)\right]}\left(\frac{u_1^2}{\sigma_{u_1}^2}+\frac{u_2^2}{\sigma_{u_2}^2}-\frac{2\,u_1u_2\rho\left(u_1,u_2\right)}{\sigma_{u_1}\sigma_{u_2}}\right)\right] = \text{constant}$$

At the points $\left(u_1,u_2\right) = \left(\pm\sigma_{u_1},0\right)$ and/or $\left(u_1,u_2\right) = \left(0,\pm\sigma_{u_2}\right)$ on the ellipse curve in the 2-D $u_1 - u_2$ plane, we see that constant = 1/2 (true for any value of $\left(u_1,u_2\right)$ lying on this particular ellipse) and thus we see that the equation of an ellipse for which $\left[\frac{1}{2}\frac{1}{\left[1-\rho^2\left(u_1,u_2\right)\right]}\left(\frac{u_1^2}{\sigma_{u_1}^2}+\frac{u_2^2}{\sigma_{u_2}^2}-\frac{2\,u_1u_2\rho\left(u_1,u_2\right)}{\sigma_{u_1}\sigma_{u_2}}\right)\right] = \frac{1}{2}$ associated with the above 2-D Gaussian/normal P.D.F. $G\left(u_1,u_2\right)$, corresponds to a 1-sigma (= 1 standard deviation) contour (*aka* 1-sigma "equipotential") of constant/equal probability density for this distribution.

Similarly, e.g. for an arbitrary # of (integer) sigma/standard deviations, i.e. $n_\sigma = 1,2,3,4,5,...$ this corresponds to contours associated with equation of ellipses for which $\left[\frac{1}{2}\frac{1}{\left[1-\rho^2\left(u_1,u_2\right)\right]}\left(\frac{u_1^2}{\sigma_{u_1}^2}+\frac{u_2^2}{\sigma_{u_2}^2}-\frac{2\,u_1u_2\rho\left(u_1,u_2\right)}{\sigma_{u_1}\sigma_{u_2}}\right)\right] = \frac{n_\sigma^2}{2}$ is satisfied, as shown in the figure below:



Note from the above discussion(s), for k ***independent*** random variables that we can also write the Gaussian P.D.F. in matrix notation as:

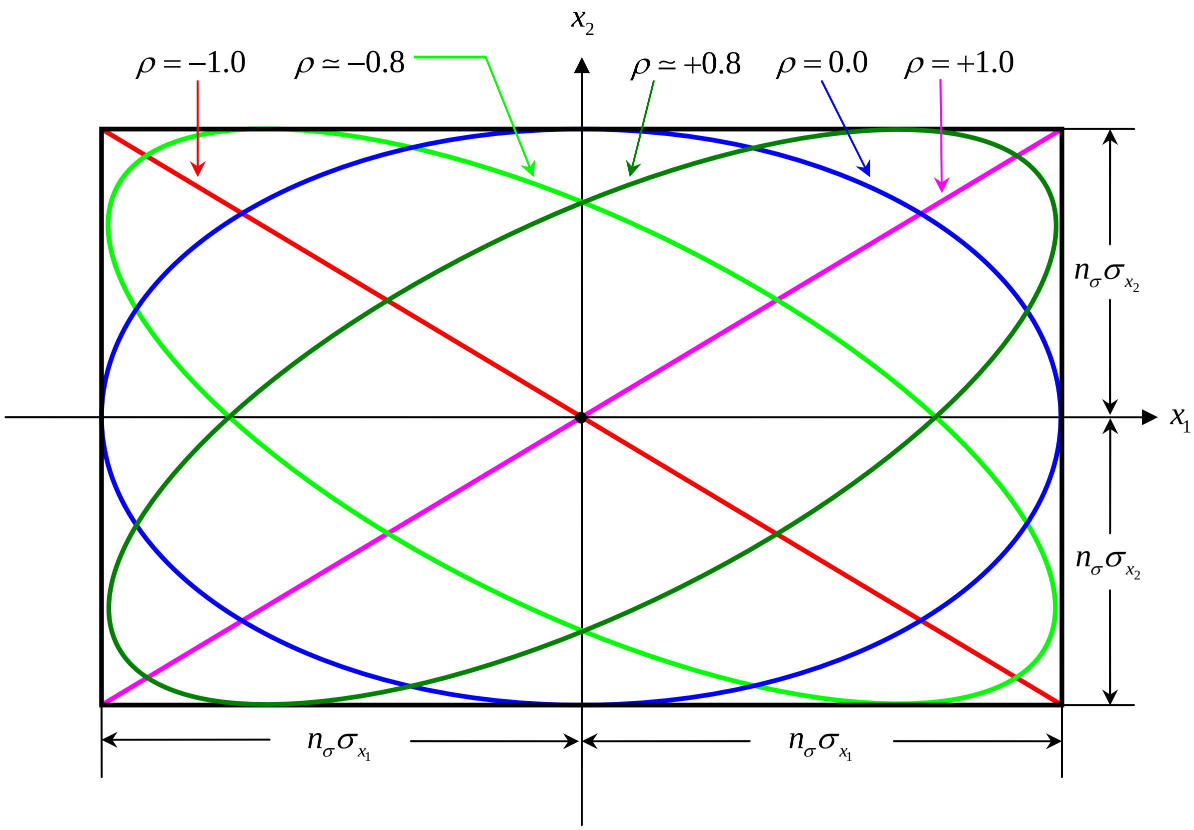
$$G(x_1, x_2, \ldots x_k) = G(x_1)\cdot G(x_2)\ldots G(x_k) = \frac{1}{(2\pi)^{k/2}\prod_{i=1}^{k}\sigma_{x_i}} e^{-\frac{1}{2}\sum_{i=1}^{k}\frac{x_i^2}{\sigma_{x_i}^2}} = \frac{1}{(2\pi)^{k/2}\left|\underline{V}_x\right|^{1/2}} e^{-\frac{1}{2}\underline{x}^T\hat{\underline{V}}_x^{-1}\underline{x}}$$

For k ***dependent*** random variables, the Gaussian P.D.F. in matrix notation is:

$$G(u_1, u_2, \ldots u_k) = \frac{1}{(2\pi)^{k/2}\prod_{i,j=1}^{k}\sigma_{u_i}\sigma_{u_j}\sqrt{1-\rho^2(u_i,u_j)}} e^{-\frac{1}{2}\sum_{i,j=1}^{k}\left[\frac{1}{\left[1-\rho^2(u_i,u_j)\right]}\left(\frac{u_i^2}{\sigma_{u_i}^2}+\frac{u_j^2}{\sigma_{u_j}^2}-\frac{2u_iu_j\rho(u_i,u_j)}{\sigma_{u_i}\sigma_{u_j}}\right)\right]} = \frac{1}{(2\pi)^{k/2}\left|\underline{V}_u\right|^{1/2}} e^{-\frac{1}{2}\underline{u}^T\hat{\underline{V}}_u^{-1}\underline{u}}$$

In the figure below, for a given n_σ ellipse contour, we show the effect of varying the correlation coefficient $-1 \le \rho(x_1, x_2) \equiv \dfrac{\operatorname{cov}(x, x_2)}{\sigma_{u_1}\sigma_{u_2}} \le +1$ between its upper/lower limits in the general n_σ ellipse equation: $\left[\dfrac{1}{2}\dfrac{1}{\left[1-\rho^2(u_1,u_2)\right]}\left(\dfrac{u_1^2}{\sigma_{u_1}^2}+\dfrac{u_2^2}{\sigma_{u_2}^2}-\dfrac{2u_1u_2\rho(u_1,u_2)}{\sigma_{u_1}\sigma_{u_2}}\right)\right] = \dfrac{n_\sigma^2}{2}$

e.g. for the $G(x_1, x_2)$ P.D.F., when x_1 and x_2 in general are ***dependent*** random variables:



Note that ***each*** of the n_σ ellipses in the above figure touch/are tangent to the sides of an enclosing rectangular box of dimensions $2n_\sigma\sigma_{x_1} \times 2n_\sigma\sigma_{x_2}$. The above n_σ ellipse equation(s), used in conjunction with the 4 straight-line equations that describe each of the four sides of the enclosing rectangular box can be solved simultaneously to determine the tangent/intersection point(s) of a given ellipse with the enclosing rectangular box.