Change of Variable

Principle Axes Theorem

If A is a symmetric matrix then there exists an orthogonal change of variable $\vec{x} = P\vec{y}$ that transforms $\vec{x}^T A\vec{x}$ to $\vec{y}^T D\vec{y}$ with no cross-product terms.

Proof

Given $Q = \vec{x}^T A\vec{x}$, where $\vec{x} \in \mathbb{R}^n$ is a variable vector and A is a real $n \times n$ symmetric matrix. Then,

$$A = PDP^T$$

where P is an $n \times n$ orthogonal matrix. A **change of variable** can be represented as

$$\vec{x} = P\vec{y}, \quad \text{or} \quad \vec{y} = P^{-1}\vec{x}$$

With this change of variable, the quadratic form $\vec{x}^T A\vec{x}$ becomes:

$$\begin{aligned} Q = \vec{x}^T A\vec{x} &= (P\vec{y})^T A(P\vec{y}) \\ &= \vec{y}^T P^T AP\vec{y} \\ &= \vec{y}^T D\vec{y}, \quad \text{using } A = PDP^T \end{aligned}$$

To express this (the theorem) more clearly, what we are doing is changing variables from $x_1, x_2, \ldots, x_n$ to $y_1, y_2, \ldots, y_n$. We will define this using equations discussed later. Our motivation to do this is to remove the cross product term, so that equation is easier to understand.

Here is the equation relating $\vec{x}$ and $\vec{y}$ that I promised earlier,

$$\vec{x} = P\vec{y} \qquad\qquad \vec{y} = P^{-1}\vec{x}$$

Now remember that we are apply this to matrix Q which is the quadratic form of an equation. Also recall that $Q = \vec{x}^T A\vec{x}$, where matrix A is symmetric. Because of this property of A we can create an orthogonal diagonalization of A where $A = PDP^T$, D is diagonal, and P is an orthonormal matrix, so $P^{-1} = P^T$ Now we can do some math,

$$
\begin{aligned}
Q &= \vec{x}^T A \vec{x} \\
Q &= \vec{x}^T P D P^T \vec{x} \\
Q &= \vec{x}^T P D \vec{y} \\
Q &= \vec{x}^T (P^T)^T D \vec{y} \\
Q &= (P^T \vec{x})^T D \vec{y} \\
Q &= (P^{-1} \vec{x})^T D \vec{y} \\
Q &= (\vec{y})^T D \vec{y} \\
Q &= \vec{y}^T D \vec{y}
\end{aligned}
$$

Now we have successful expressed Q in terms of $\vec{y}$ instead of $\vec{x}$, and more importantly the middle matrix is D. Remember D is a diagonal matrix, hence $\vec{y}^T D \vec{y}$ will **not** have any no cross product terms. Don't believe me? Well just multiply any diagonal matrix D with a correspondingly sized $\vec{y}$.