Appendix F: Meaning of the Determinant of Covariance Matrix

Connection to Differential Entropy

To put it in other words: Let's say you have a (large) set of points from which you assume it is Gaussian distributed. If you compute the determinant of the sample covariance matrix then you measure (indirectly) the differential entropy of the distribution up to constant factors and a logarithm. See, e.g, Multivariate normal distribution.

The differential entropy of a Gaussian density is defined as:

$$H[p] = \frac{k}{2}(1 + \ln(2\pi)) + \frac{1}{2}\ln|\Sigma|$$

where k is the dimensionality of your space, i.e., in our case k = 3

Meaning of the Determinant of Covariance Matrix

is positive semi-definite, which means ||0. **The larger ||, the more are your data points dispersed.**
If ||=0, it means that your data ponts do not 'occupy the whole space', **meaning that they lie, e.g., on a line or a plane within 3.**
Somewhere I have read, that || is also called generalized variance. It captures the volume of your data cloud.

Since a sample covariance matrix is defined somewhat like:

$$\Sigma = \frac{1}{N-1}\sum_{i=1}^{N}(\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T$$

it follows, that you do not capture any information about the mean. You can verify that easily by adding some large constant vectorial shift to your data; || should not change.

Connection to PCA

the eigenvalues 1,2,3 of correspond to the variances along the principal component axis of your data points,
|| captures their product, because by definition the determinant of a matrix is equal to the product of its eigenvalues.

Note that the largest eigenvalue corresponds to the maximal variance w.r.t. to your data (direction given by the corresponding eigenvector, see PCA).