multivariate normal distribution

- The vector $\vec{X}$ is a standard normal collection if each $\vec{X}_i$ is a mutually independent standard normal.
- The random vector $(X_1..X_n)$ is multivariate normal if every linear combination $t_1X_1 + ..t_nX_n$ is normal.
- A covariance matrix (also known as dispersion matrix or variance–covariance matrix) is a matrix whose element in the *i*, *j* position is the covariance between the *i*th and *j*th elements of a random vector. A random vector is a random variable with multiple dimensions. Each element of the vector is a scalar random variable. When working with multiple variables, the covariance matrix provides a succinct way to summarize the covariances of all pairs of variables.
- Because the random variables are independent, the covariance matrix (which is a matrix with the covariance as elements) is the identity matrix. (Remember that thing about independent variables having 0 covariance? → from the Stat 110 class)
- It can be easily shown that
  $\Sigma = E[(X-\mu)(X-\mu)^T] = E[XX^T] - \mu\mu^T$
  this just follows from the definitions of covariance matrix and basic rules of matrix multiplication and expectation.
- Given the definition, that a multivariate normal is just a vector $\vec{X}$ so that all $t_1X_1 + ... + t_nX_n$are normal, we can derive with some manipulation that the multivariate pdf is given by:

$$p(x;\mu,\Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right).$$

(Note that as we take different linear combinations $t_1X_1 + ...t_nX_n$, and put the results in a vector, that covariance matrix is not, in general, digonal anymore).

- Note that Σ is always symmetric. It can be easily shown that given any covariance matrix and for any $\vec{x}$ in R^n, $\vec{x}\,\Sigma\,x \geq 0$.
- Because the pdf above is just a regular pdf over multiple variables (kind of like joint distribution with some covariances that are non-0), it should still sum upto 1:

$$\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)dx_1dx_2\cdots dx_n = 1.$$

- Another way to understand a multivariate Gaussian conceptually is to understand the shape of its isocontours, which is the set of all $\vec{x}$ that give the same value of the pdf.
- Note that by our definition of the multivariate normal, any random variable X with a multivariate Gaussian distribution can be interpreted as the result of applying a linear transformation (X = BZ + μ) to some collection of n independent standard normal random variables (Z).
- proof idea for sum of gaussians is a gaussian:

- The proof uses the fact that the moment generating function (mgf) of the sum of independent random variables is the product of the respective moment generating functions. After computing the mgf of a normal and taking the product of two mgfs, we see that the product is again the mgf of a normal random variable. Then the proof follows by using the uniqueness theorem for an mgf, that is, the fact that the moment generating function is uniquely determined by the distribution.