Isotropy

Remark. For a random vector $X = (X_1, \ldots, X_n)$, recall that

$$\mathrm{cov}(X) = \mathbb{E}\, X \cdot X^T - \mu \cdot \mu^T$$

$$= \Sigma - \mu \cdot \mu^T$$

The (i,j)-th entries of $\mathrm{cov}(X)$ and Σ are

$$\mathrm{cov}(X)_{ij} = \mathbb{E}\,(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)$$

[covariance of X_i and X_j]

and

$$\Sigma_{ij} = \mathbb{E}\, X_i X_j .$$

Def. A random vector $X = (X_1, \ldots, X_n)$ is called **isotropic** if

$$\Sigma(X) = \mathbb{E}\, X X^T = I_n ,$$

where I_n is the $n \times n$ identity matrix.

Isotropy is the higher-dimensional analogue of unit variance (if mean zero) or $\mathbb{E}X^2 = 1$.

For random variables, it's common to center and scale to get the **standard score** (or **z-score**)

$$Z = \frac{X - \mathbb{E}X}{\sqrt{\mathrm{Var}(X)}}. \quad \begin{pmatrix} \text{mean zero} \\ \text{unit variance} \end{pmatrix}$$

We can do something similar for random vectors:

Reduction to Isotropy

Suppose X has mean μ and $\mathrm{cov}(X) = \Sigma$ is invertible. Then

$$Z = \Sigma^{-1/2}(X - \mu)$$

is isotropic with mean zero.

Characterizing Isotropy

Lemma 3.2.3 A random vector X in $\mathbb{R}^n$ is isotropic **if and only if**

$$\mathbb{E}[\langle X, x\rangle^2] = \|x\|_2^2$$

for all $x \in \mathbb{R}^n$.

Proof. Two real symmetric matrices, A and B, are equal **if and only if**

$$x^T A x = x^T B x$$

for all $x \in \mathbb{R}^n$. Why?

[**Hint:** consider $x^T(A-B)x = 0$ for certain choices of x and use symmetry.]

Recall that X is isotropic when $\mathbb{E} X X^T = I_n$.
Hence X is isotropic **if and only if**

$$x^T(\mathbb{E} X X^T)x = x^T \cdot I_n \cdot x \text{ for all } x \in \mathbb{R}^n.$$

The LHS is

$$\mathbb{E}(x^T X X^T x) = \mathbb{E}\langle x, X\rangle\langle X, x\rangle$$
$$= \mathbb{E}\langle X, x\rangle^2$$

On the other hand, the RHS is

$$x^T \cdot I_n \cdot x = x^T \cdot x = \langle x, x\rangle = \|x\|_2^2. \quad \square$$

Equivalently, X is isotropic **if and only if**

$$\mathbb{E}\langle X, u\rangle^2 = 1$$

for all unit vectors $u \in \mathbb{R}^n$.

[a one-dimensional marginal of X]

Recall that $\langle X, u\rangle$ is the projection of X along u. Hence, the lemma says that isotropic distributions tend to extend evenly in all directions.

Lemma 3.2.4 Let X (in $\mathbb{R}^n$) be isotropic. Then

$$\mathbb{E}\,\|X\|_2^2 = n.$$

If X and Y are independent and isotropic, then

$$\mathbb{E}\langle X, Y\rangle^2 = n.$$

Proof. We have

$$\mathbb{E}\,\|X\|_2^2 = \mathbb{E}\langle X, X\rangle$$

$$\left(= \mathbb{E}\, X^T \cdot X\right)$$

trace

$$= \mathbb{E}\,\mathrm{tr}(X^T \cdot X)$$

[$\mathrm{tr}(AB) = \mathrm{tr}(BA)$] $\qquad = \mathbb{E}\,\mathrm{tr}(X \cdot X^T)$

[linearity of tr] $\qquad = \mathrm{tr}(\mathbb{E}\, X \cdot X^T).$

Since X is isotropic, this is $\mathrm{tr}(I_n) = n$.

For the second part, we condition on Y. The **law of total expectation** says that

$$\mathbb{E}\langle X,Y\rangle^2 = \mathbb{E}_Y\, \mathbb{E}_X\left[\langle X,Y\rangle^2 \mid Y\right].$$

$$\text{[Lemma 3.2.3]} \quad = \mathbb{E}_Y \|Y\|_2^2$$

$$\text{[first part of proof]} \quad = n. \qquad \square$$

Remark. We've now seen that

$$\|X\|_2^2,\ \|Y\|_2^2 \approx n$$

and also

$$\sqrt{n} \approx |\langle X,Y\rangle|$$

$$= \|X\|_2 \cdot \|Y\|_2 \cdot |\cos\theta| \quad \text{[angle between } X, Y\text{]}$$

$$\approx n \cdot |\cos\theta|.$$

Then $|\cos\theta| \approx \frac{1}{\sqrt{n}}$, which is ≈ 0 for large n. Hence $\theta \approx \pm\frac{\pi}{2}$. That is, independent isotropic random vectors tend to be nearly orthogonal!

Examples of Isotropic Distributions

1) If $X_1, X_2, \ldots, X_n$ are independent random variables with mean zero and unit variance, then $X = (X_1, \ldots, X_n)$ is isotropic. Why? The entries of Σ are

$$\mathbb{E}\, X_i X_j = \begin{cases} \mathbb{E}\, X_i^2 & \text{if } i = j, \\ \mathbb{E}\, X_i X_j & \text{if } i \neq j, \end{cases}$$

$$= \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Thus $\Sigma = I_n$, so X is isotropic.

2) (special case) If each X_i in 1) is a symmetric Bernoulli random variable [so $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$], then X is a symmetric Bernoulli random vector. Equivalently, X is the uniform distribution on the unit discrete cube $\{-1, 1\}^n$ in $\mathbb{R}^n$.

Coordinates need not be independent, however.

3) Suppose X is uniformly distributed on the sphere of radius $\sqrt{n}$ centered at the origin in $\mathbb{R}^n$. If $X = (X_1, \ldots, X_n)$, then

$$X_1^2 + \cdots + X_n^2 = n.$$

Thus $$X_n = \pm\sqrt{X_1^2 + \cdots + X_{n-1}},$$

which depends on the other coordinates.
[Also called the **spherical distribution** in $\mathbb{R}^n$.]

Multivariate Gaussians

A random vector $g = (g_1, \ldots, g_n)$ is said to have the **standard normal** distribution, written $g \sim N(0, I_n)$, if $g_1, \ldots, g_n \sim N(0,1)$ are independent.

Since the coordinates are independent with mean zero and unit variance, g is isotropic.

Its PDF is given by

$$f(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-x_i^2/2}$$

$$= \frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2}, \quad x \in \mathbb{R}^n.$$

Note. The density only relies on the length of x, so it is **rotation invariant**.

Prop. Suppose U is an orthogonal $n \times n$ matrix, and $g \sim N(0, I_n)$. Then

$$Ug \sim N(0, I_n).$$

We get a general normal distribution as follows. Let $\mu \in \mathbb{R}^n$ and suppose Σ is an invertible, positive-semidefinite matrix. If $Z \sim N(0, I_n)$, then

$$X := \mu + \Sigma^{1/2} \cdot Z$$

is **normally distributed** with mean μ and covariance matrix Σ (check!). Its density is given by

$$f_X(x) = \frac{1}{(2\pi)^{n/2} \cdot (\det \Sigma)^{1/2}} e^{-\frac{1}{2}(x-\mu)^T \cdot \Sigma \cdot (x-\mu)}$$

for $x \in \mathbb{R}^n$.

Remark. If $X \sim N(\mu, \Sigma)$, then the coordinates of X are independent **if and only if** they are uncorrelated! In such case, it follows that $\Sigma = I_n$ (check!).

[Generally, independence implies uncorrelated, but not the other way around.]

Also, we know from **Theorem 3.1.1** that $g \sim N(0, I_n)$ concentrates around the sphere of radius $\sqrt{n}$. Thus, in high dimensions, the standard normal is close to the uniform distribution on the sphere. This behavior differs greatly from the standard normal in $\mathbb{R}$!