

Gaussian random vectors

- ▶ The Gamma function
- ▶ The χ^2 distribution
- ▶ Confidence ellipsoids
- ▶ Marginal density functions
- ▶ Degenerate Gaussian random vectors
- ▶ Changes of variables for random vectors
- ▶ Linear transformations of Gaussians

Random variables

- ▶ we have a *real-valued random variable* x with *probability density function* (pdf) so that

$$\text{Prob}(x \in [a, b]) = \int_a^b p(x) dx$$

- ▶ the *mean* or *expected value* of x is $\mathbf{E}(x) = \int_{-\infty}^{\infty} xp(x) dx$

- ▶ the *variance* of x is $\mathbf{var}(x) = \mathbf{E}((x - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx$

- ▶ mean is linear, so for $a, b \in \mathbb{R}$

- ▶ $\mathbf{E}(ax + b) = a \mathbf{E}(x) + b$

- ▶ $\mathbf{var}(ax + b) = a^2 \mathbf{var}(x)$

Gaussian random variables

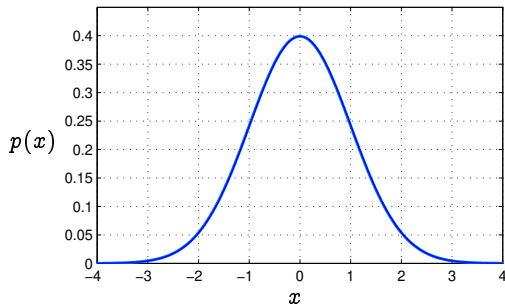
- ▶ x is *Gaussian* if it has *probability density function* (pdf) given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- ▶ write this as $x \sim \mathcal{N}(\mu, \sigma^2)$
- ▶ the *mean* or *expected value* of x is $\mathbf{E}(x) = \mu$
- ▶ the *variance* of x is $\mathbf{E}((x - \mu)^2) = \sigma^2$

Gaussian random variables

pdf for $x \sim \mathcal{N}(0, 1)$ is



- ▶ p is symmetric about the mean
- ▶ decays very fast; but $p(x) > 0$ for all x

Computing probabilities for Gaussian random variables

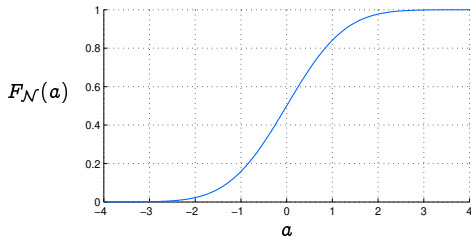
- ▶ the *error function* is

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

- ▶ the Gaussian *cumulative distribution function* (CDF) is

$$F_{\mathcal{N}}(a) = \mathbf{Prob}(x \leq a) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{a - \mu}{\sigma\sqrt{2}}\right)$$

- ▶ plot shows $\mu = 0$ and $\sigma = 1$,



Computing probabilities for Gaussian random variables

- ▶ for $x \sim \mathcal{N}(0, \sigma^2)$ we have for $a \geq 0$

$$\mathbf{Prob}(x \in [-a, a]) = \mathbf{erf}\left(\frac{a}{\sigma\sqrt{2}}\right)$$

- ▶ some particular values:

- ▶ $\mathbf{Prob}(x \in [-\sigma, \sigma]) \approx 0.68$
- ▶ $\mathbf{Prob}(x \in [-2\sigma, 2\sigma]) \approx 0.9545$
- ▶ $\mathbf{Prob}(x \in [-3\sigma, 3\sigma]) \approx 0.9973$

Continuous random vectors

- ▶ suppose \mathbb{R}^n -valued random vector x has probability density function $p^x : \mathbb{R}^n \rightarrow \mathbb{R}$.
- ▶ the *mean* or *expected value* of x is

$$\mathbf{E}(x) = \int_{\mathbb{R}^n} x p^x(x) dx$$

Covariance

- ▶ the *covariance* of x is

$$\mathbf{cov}(x) = \mathbf{E}((x - \mu)(x - \mu)^T) = \int_{\mathbb{R}^n} (x - \mu)(x - \mu)^T p^x(x) dx$$

- ▶ We'll often denote the covariance by $\Sigma = \mathbf{cov}(x)$
- ▶ Σ is symmetric and positive semidefinite
- ▶ Σ_{ii} is the covariance of the i 'th component x_i

$$\Sigma_{ii} = \mathbf{cov}(x_i)$$

Affine transformations

- ▶ suppose x is an \mathbb{R}^n -valued random vector
- ▶ let y be an *affine function* of x , given by $y = Ax + b$
- ▶ the mean of y is the same affine function of the mean of x

$$\mathbf{E}(y) = A \mathbf{E}(x) + b$$

- ▶ the covariance of y is a linear function of the covariance of x

$$\mathbf{cov}(y) = A \mathbf{cov}(x) A^T$$

Affine transformations of random vectors

- for the mean, we have

$$\begin{aligned}\mathbf{E} y &= \mathbf{E}(Ax + b) \\ &= \int_{\mathbb{R}^n} (Ax + b)p(x) dx \\ &= b + A \int_{\mathbb{R}^n} xp(x) dx \\ &= b + A \mathbf{E} x\end{aligned}$$

- and for the covariance

$$\begin{aligned}\mathbf{cov}(y) &= \mathbf{E}((y - \mathbf{E} y)(y - \mathbf{E} y)^T) \\ &= \mathbf{E}(A(x - \mathbf{E} x)(x - \mathbf{E} x)^T A^T) \\ &= A \mathbf{E}((x - \mathbf{E} x)(x - \mathbf{E} x)^T) A^T \\ &= A \mathbf{cov}(x) A^T\end{aligned}$$

Mean-square deviation

- ▶ Suppose x is an \mathbb{R}^n -valued random variable, with mean μ .
- ▶ The *mean square deviation from the mean* is given by

$$\mathbf{E}(\|x - \mu\|^2) = \mathbf{trace\,cov}(x)$$

- ▶ because

$$\begin{aligned}\mathbf{E}(\|x - \mu\|^2) &= \mathbf{E}((x - \mu)^T(x - \mu)) \\ &= \mathbf{E\,trace}((x - \mu)^T(x - \mu)) \\ &= \mathbf{E\,trace}((x - \mu)(x - \mu)^T) && \text{since } \mathbf{trace}(AB) = \mathbf{trace}(BA) \\ &= \mathbf{trace\,E}((x - \mu)(x - \mu)^T) && \text{since } \mathbf{E\,Ax} = A\mathbf{E\,x}\end{aligned}$$

The mean-variance decomposition

- ▶ the *mean square* of a random variable x is

$$\mathbf{E}(\|x\|^2) = \text{trace}(\text{cov}(x)) + \|\mathbf{E} x\|^2$$

- ▶ this holds because

$$\begin{aligned}\mathbf{E}(\|x\|^2) &= \mathbf{E}(\|x - \mu + \mu\|^2) \\ &= \mathbf{E}(\|x - \mu\|^2 + 2\mu^T(x - \mu) + \|\mu\|^2) \\ &= \mathbf{E}(\|x - \mu\|^2) + 2\mu^T \mathbf{E}(x - \mu) + \|\mu\|^2\end{aligned}$$

- ▶ called the *mean-variance decomposition*

Correlation coefficient

- ▶ let $\Sigma = \text{cov}(x)$. The *correlation coefficient* of x_i and x_j is

$$\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$$

- ▶ Since $\Sigma \succeq 0$, we have $|\rho_{ij}| \leq 1$
- ▶ If $\rho_{ij} = 0$ then x_i and x_j are called *uncorrelated*.

Correlation and covariance

- ▶ the *correlation matrix* of random vector x is

$$\mathbf{corr}(x) = \mathbf{E}(xx^T)$$

- ▶ not to be confused with the correlation coefficient!
- ▶ If $\mathbf{E} x = 0$ then $\mathbf{corr}(x) = \mathbf{cov}(x)$
- ▶ The *mean square* of x is $\mathbf{E}(\|x\|^2) = \text{trace } \mathbf{corr}(x)$
- ▶ The *correlation-covariance decomposition* is

$$\mathbf{corr}(x) = \mathbf{cov}(x) + (\mathbf{E} x)(\mathbf{E} x^T)$$

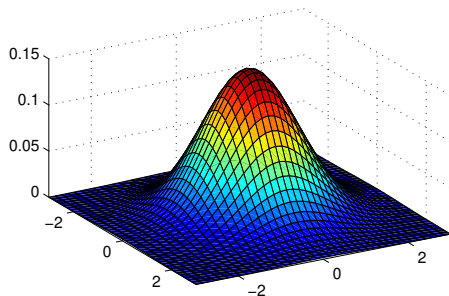
- ▶ same approach as the mean-variance formula

Gaussian random vectors

- ▶ the \mathbb{R}^n -valued random variable x is called *Gaussian* if it has pdf

$$p^x(x) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right)$$

- ▶ write this as $x \sim \mathcal{N}(\mu, \Sigma)$, here $\Sigma = \Sigma^T$ and $\Sigma > 0$



Gaussian random vectors

- ▶ suppose $x \sim \mathcal{N}(\mu, \Sigma)$. Then the mean of x is

$$\mathbf{E} x = \mu$$

- ▶ and the covariance of x is

$$\mathbf{cov}(x) = \Sigma$$

Ellipsoids

- ▶ the Gaussian pdf is constant on the surface of the ellipsoids

$$S_\alpha = \left\{ x \in \mathbb{R}^n \mid (x - \mu)^T \Sigma^{-1} (x - \mu) \leq \alpha \right\}$$

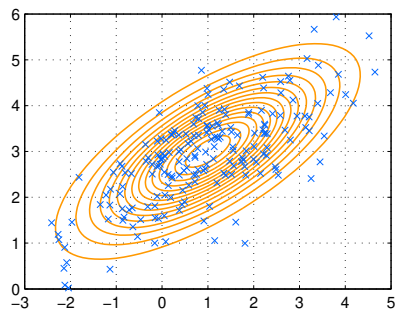
- ▶ center is at μ , semiaxis lengths are $\sqrt{\alpha \lambda_i(\Sigma)}$.

- ▶ example has

$$\mu = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

contours at $p(x) = 0.01, 0.02, \dots$



Gamma function

- ▶ the *gamma function* is

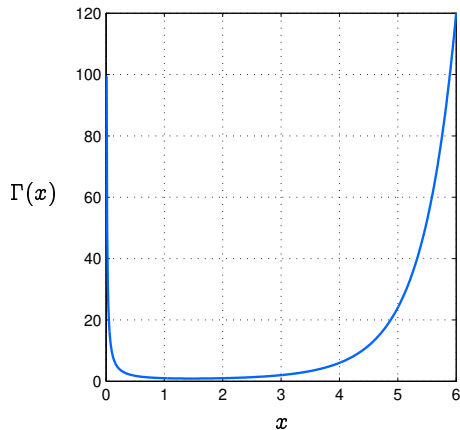
$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad \text{for } x > 0$$

- ▶ for $x > 0$

$$\Gamma(x+1) = x\Gamma(x)$$

- ▶ $\Gamma(1) = 1$, so for integer $x > 1$

$$\Gamma(x) = (x-1)!$$

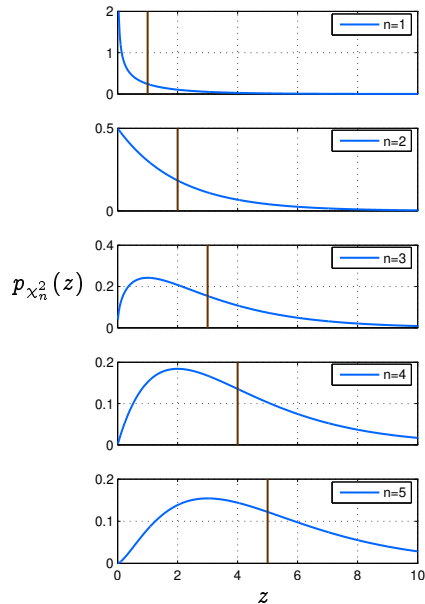


The χ^2 distribution

- ▶ the χ_n^2 probability density function is

$$p_{\chi_n^2}(z) = \frac{1}{2^{\frac{n}{2}} \Gamma(n/2)} z^{\frac{n}{2}-1} e^{-\frac{z}{2}}$$

- ▶ A family of pdfs, one for each $n > 0$
- ▶ If $z \sim \chi_n^2$, then $\mathbf{E} z = n$



Gaussian random vectors and confidence ellipsoids

- ▶ suppose x is Gaussian, i.e., $x \sim \mathcal{N}(\mu, \Sigma)$, where $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n \times n}$. Define the random variable

$$z = (x - \mu)^T \Sigma^{-1} (x - \mu)$$

which is a measure of the distance of x from μ

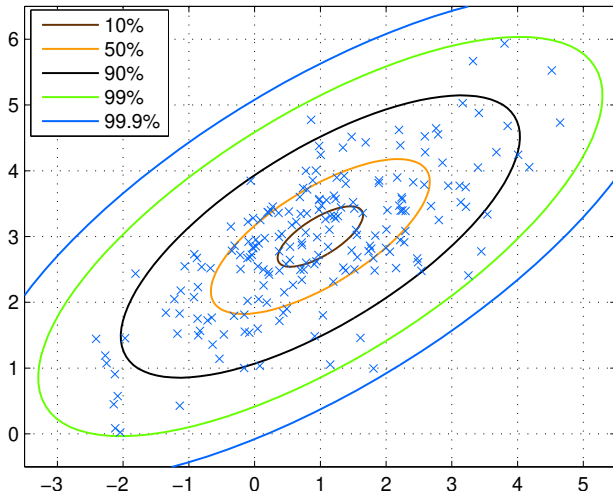
- ▶ z has a χ_n^2 distribution
- ▶ Hence prob. that x lies in the ellipsoid $S_\alpha = \{ x \in \mathbb{R}^n \mid (x - \mu)^T \Sigma^{-1} (x - \mu) \leq \alpha \}$

$$\text{Prob}(x \in S_\alpha) = F_{\chi_n^2}(\alpha)$$

- ▶ for example $F_{\chi_n^2}(\alpha) \approx \begin{cases} \frac{1}{2} & \text{if } \alpha = n \\ 0.9 & \text{if } \alpha = n + 2\sqrt{n} \end{cases}$ 90% confidence ellipsoid

Confidence ellipsoids

The plot shows the confidence ellipsoids and 200 sample points.



Marginal probability density functions

- ▶ suppose x is an RV with pdf $p^x : \mathbb{R}^n \rightarrow \mathbb{R}$, and $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, where $x_1 \in \mathbb{R}^r$.
- ▶ define the *marginal pdf* of x_1 to be the function p^{x_1} such that

$$\mathbf{Prob}(x_1 \in W) = \int_W p^{x_1}(z) dz \quad \text{for all } W \subset \mathbb{R}^r$$

- ▶ we also know that

$$\mathbf{Prob}(x_1 \in W) = \int_W \int_{x_2 \in \mathbb{R}^{n-r}} p^x(x_1, x_2) dx_2 dx_1$$

- ▶ since these are equal, we have

$$p^{x_1}(x_1) = \int_{x_2 \in \mathbb{R}^{n-r}} p^x(x_1, x_2) dx_2$$

The marginal pdf of a Gaussian

- ▶ suppose $x \sim \mathcal{N}(\mu, \Sigma)$, and

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

- ▶ let's look at the component x_1
- ▶ Since $x_1 = \begin{bmatrix} I & 0 \end{bmatrix} x$, we have the mean

$$\mathbf{E} x_1 = \begin{bmatrix} I & 0 \end{bmatrix} \mu = \mu_1$$

and also the covariance

$$\mathbf{cov}(x_1) = \begin{bmatrix} I & 0 \end{bmatrix} \Sigma \begin{bmatrix} I \\ 0 \end{bmatrix} = \Sigma_{11}$$

- ▶ In fact, the random variable x_1 is *Gaussian*; this is not obvious

Proof: the marginal pdf of a Gaussian

- assume for convenience that $\mathbf{E} x = 0$. The marginal pdf of x_1 is

$$p^{x_1}(x_1) = \int_{x_2} c_1 \exp\left(-\frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) dx_2$$

- we have, by the completion of squares formula

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}^{-1} = \begin{bmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & (\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{bmatrix}$$

- and so, setting $S = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^T \Sigma_{11}^{-1} x_1 + (x_2 - \Sigma_{21}\Sigma_{11}^{-1}x_1)^T S^{-1} (x_2 - \Sigma_{21}\Sigma_{11}^{-1}x_1)$$

Proof: the marginal pdf of a Gaussian

► hence we have

$$\begin{aligned} p^{x_1}(x_1) &= c_1 \exp\left(-\frac{1}{2}x_1^T \Sigma_{11}^{-1} x_1\right) \int_{x_2} \exp\left(-\frac{1}{2}(x_2 - \Sigma_{21}\Sigma_{11}^{-1}x_1)^T S^{-1}(x_2 - \Sigma_{21}\Sigma_{11}^{-1}x_1)\right) dx_2 \\ &= c_2 \exp\left(-\frac{1}{2}x_1^T \Sigma_{11}^{-1} x_1\right) \end{aligned}$$

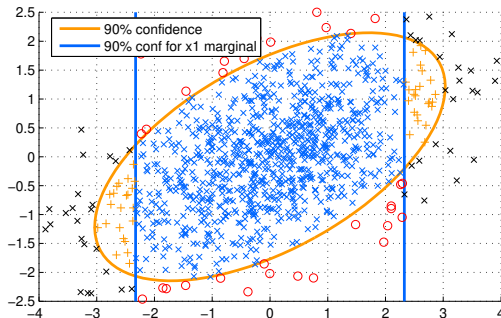
► now c_2 is determined, because $\int p^{x_1}(z) dz = 1$, so we don't need to calculate it explicitly.

► therefore, if $x \sim \mathcal{N}(0, \Sigma)$ the marginal pdf of x_1 is *Gaussian*, and

$$x_1 \sim \mathcal{N}(0, \Sigma_{11})$$

Example: marginal pdf for Gaussians

- ▶ Suppose $\Sigma = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 1 \end{bmatrix}$ and $x \sim \mathcal{N}(0, \Sigma)$. A simulation of 1000 points is below
- ▶ all blue and orange points (908) are within 90% confidence ellipsoid for x
- ▶ all blue and red points (899) are within 90% confidence interval for x_1



Degenerate Gaussian random vectors

- ▶ it's convenient to allow Σ singular, but still $\Sigma = \Sigma^T$ and $\Sigma \geq 0$
this means that in some directions, *x is not random at all*
- ▶ obviously density formula does not hold; instead write

$$\Sigma = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}^T$$

where $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ is orthogonal, and $\Sigma_1 > 0$

columns of Q_1 are orthonormal basis for **range**(Σ)

columns of Q_2 are orthonormal basis for **null**(Σ)

- ▶ let $\begin{bmatrix} z \\ w \end{bmatrix} = Q^T x$; then
 - ▶ $z \sim \mathcal{N}(Q_1^T \mu, \Sigma_1)$ is non-degenerate Gaussian
 - ▶ $w = Q_2^T \mu$ is not random

Changes of variables for random vectors

- ▶ suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous, and $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfies
 - ▶ h is one-to-one and onto; i.e., h is invertible
 - ▶ Both h and h^{-1} are differentiable, with continuous derivative

- ▶ the derivative of h at x is $Dh(x)$, the *Jacobian* matrix

$$(Dh(x))_{ij} = \frac{\partial h_i}{\partial x_j}(x)$$

- ▶ then for any $A \subset \mathbb{R}^n$

$$\int_{h(A)} f(x) dx = \int_A f(h(y)) |\det Dh(y)| dy$$

Changes of variables for random vectors

- suppose x is an \mathbb{R}^n -valued random vector, and $y = g(x)$, where g is invertible, and g and g^{-1} are continuously differentiable. Then

$$p^y(y) = \frac{p^x(g^{-1}(y))}{|\det(Dg)(g^{-1}(y))|}$$

- this holds because

$$\begin{aligned}\mathbf{Prob}(y \in A) &= \int_A p^y(y) dy \\ &= \int_{g^{-1}(A)} p^x(x) dx \\ &= \int_A \frac{p^x(g^{-1}(y))}{|\det(Dg)(g^{-1}(y))|} dy\end{aligned}$$

where $D(g^{-1})(y) = \left((Dg)(g^{-1}(y)) \right)^{-1}$

Example: linear transformations

- consider $y = Ax + b$, where $A \in \mathbb{R}^{n \times n}$ is *invertible*. Then

$$p^y(y) = \frac{p^x(A^{-1}(y - b))}{|\det A|}$$

Linear transformations of Gaussians

- ▶ fundamental result: a linear function of a Gaussian random vector is a Gaussian random vector
- ▶ suppose $x \sim \mathcal{N}(\mu_x, \Sigma_x)$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Consider the linear function of x

$$y = Ax + b$$

- ▶ we already know how means and covariances transform; we have

$$\mathbf{E}(y) = A \mathbf{E} x + b \qquad \mathbf{cov}(y) = A \mathbf{cov}(x) A^T$$

- ▶ additional fact is that y is *Gaussian*

Linear transformations of Gaussians

► to show this, first suppose $A \in \mathbb{R}^{n \times n}$ is *invertible*. Let $\mu_y = A\mu_x + b$ and $\Sigma_y = A\Sigma_x A^T$.

► we know

$$p^x(x) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma_x)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma_x^{-1}(x - \mu)\right)$$

► so

$$\begin{aligned} p^y(y) &= \frac{p^x(A^{-1}(y - b))}{|\det A|} \\ &= \frac{1}{|\det A|(2\pi)^{\frac{n}{2}} (\det \Sigma_x)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(y - b - A\mu_x)^T (A^{-1})^T \Sigma_x^{-1} A^{-1}(y - b - A\mu_x)\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma_y)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(y - \mu_y)^T \Sigma_y^{-1}(y - \mu_y)\right) \end{aligned}$$

Non-invertible linear transformations of Gaussians

- suppose $A \in \mathbb{R}^{m \times n}$, and $y = Ax$ where $x \sim \mathcal{N}(0, \Sigma_x)$. The SVD of A is

$$A = U\Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

- so we decompose the map into

$$y = Uw \quad \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \Sigma z \quad \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = V^T x$$

Non-invertible linear transformations of Gaussians

- ▶ since V is invertible, we know $z \sim \mathcal{N}(0, \Sigma_z)$, where

$$\Sigma_z = \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \Sigma_x \begin{bmatrix} V_1 & V_2 \end{bmatrix}$$

- ▶ we know z is Gaussian, hence the marginal z_1 is Gaussian

$$z_1 \sim \mathcal{N}(0, V_1^T \Sigma_x V_1)$$

- ▶ also $w_2 = 0$, and since Σ_1 is invertible, w_1 is Gaussian

$$w_1 \sim \mathcal{N}(0, \Sigma_1 V_1^T \Sigma_x V_1 \Sigma_1)$$

- ▶ since $w = U^T y$, we have y is a *degenerate Gaussian random vector* where

- ▶ $w_1 = U_1^T y$ are the components of y that are Gaussian
- ▶ $w_2 = 0$ are the components of y that are not random

Full-rank case

- ▶ when $\text{range}(A) = \mathbb{R}^m$, i.e., A is full row rank, we have

$$y \sim \mathcal{N}(0, A\Sigma_x A^T)$$

- ▶ because the SVD of A is

$$A = U \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

- ▶ then $y = Uw_1$, and since U is invertible, we have

$$y \sim \mathcal{N}(0, U\Sigma_1 V_1^T \Sigma_x V_1 \Sigma_1 U^T)$$

Example: simulating Gaussian random vectors

- ▶ in many languages its easy to generate $x \sim \mathcal{N}(0, I)$
- ▶ to generate $y \sim \mathcal{N}(\mu, \Sigma)$, we can use

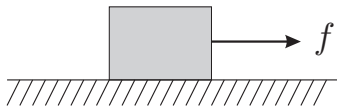
$$y = \Sigma^{\frac{1}{2}} x + \mu$$

- ▶ extremely useful for simulation

Example: Gaussian random force on mass

- ▶ x is the sequence of applied forces, so $f(t) = x_j$ for t in the interval $[j - 1, j]$.
- ▶ y_1, y_2 are final position and velocity
- ▶ $y = Ax$ where $A = \begin{bmatrix} 9.5 & 8.5 & 7.5 & 6.5 & 5.5 & 4.5 & 3.5 & 2.5 & 1.5 & 0.5 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$
- ▶ suppose the forces are Gaussian, and the vector $x \sim \mathcal{N}(0, \Sigma)$, where

$$\Sigma = \begin{bmatrix} 2 & 1 & & & & & & & & \\ & 1 & 2 & 1 & & & & & & \\ & & 1 & 2 & 1 & & & & & \\ & & & 1 & 2 & 1 & & & & \\ & & & & 1 & 2 & 1 & & & \\ & & & & & 1 & 2 & 1 & & \\ & & & & & & 1 & 2 & 1 & \\ & & & & & & & 1 & 2 & 1 \\ & & & & & & & & 1 & 2 \end{bmatrix}$$



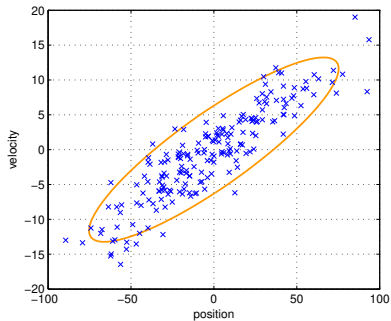
Example: Gaussian random force on mass

- ▶ the covariance of y is

$$\Sigma_y = A \Sigma A^T$$

- ▶ the 90% confidence ellipsoid is

$$\left\{ y \in \mathbb{R}^2 \mid y^T \Sigma_y^{-1} y \leq F_{\chi_n^2}^{-1}(0.9) \right\}$$



Components of a Gaussian random vector

- ▶ suppose $x \sim \mathcal{N}(0, \Sigma)$, and let $c \in \mathbb{R}^n$ be a unit vector
- ▶ let $y = c^T x$
- ▶ y is the component of x in the direction c
- ▶ y is Gaussian, with $\mathbf{E} y = 0$ and $\mathbf{cov}(y) = c^T \Sigma c$
- ▶ So $\mathbf{E}(y^2) = c^T \Sigma c$
- ▶ The unit vector c that minimizes $c^T \Sigma c$ is the eigenvector of Σ with the smallest eigenvalue. Then

$$\mathbf{E}(y^2) = \lambda_{\min}$$