STA9715 - Test 2 - Formula Sheet

Inequalities

- If $\mathsf{support}(X)$ is non-negative, $\mathbb{P}(X > x) \leq \mathbb{E}[X]/x$ (Markov)
- For any $X \sim (\mu, \sigma^2)$, $\mathbb{P}(|X-\mu| \geq k\sigma) \leq 1/k^2$ or $\mathbb{P}(|X-\mu| \geq k) \leq \sigma^2/k^2$ (Chebyshev)

Vector Arithmetic and Linear Algebra

- Vector addition $\boldsymbol{x}+\boldsymbol{y} = (x_1+y_1, x_2+y_2, \dots, x_n+y_n)$ Vector-times-scalar $\alpha\boldsymbol{x} = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$. (Both elementwise)
- Two-vector (dot / inner) product yields scalar: $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x} \cdot \boldsymbol{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$
- Vector norm: $\|\boldsymbol{x}\| = \sqrt{\sum_i x_i^2}$ - generalizes length or absolute value. Angle between vectors: $\cos\angle(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y}\rangle / \|\boldsymbol{x}\|\|\boldsymbol{y}\|$
- Matrix-vector multiplication: yields a vector: $\boldsymbol{Ax}$ element i is dot product of row i of $\boldsymbol{A}$ with $\boldsymbol{x}$.
- Matrix-matrix multiplication: yields a matrix: $\boldsymbol{AB}$ element (i,j) is dot product of row i of $\boldsymbol{A}$ with column j of $\boldsymbol{B}$.
- Quadratic form: $\langle \boldsymbol{x}, \boldsymbol{Ax}\rangle = \|\boldsymbol{x}\|_{\boldsymbol{A}}^2 = \sum_{(i,j)} A_{ij}x_ix_j$. $\boldsymbol{A}$ is *positive-definite* if all quadratic forms are positive (for $\boldsymbol{x} \neq \boldsymbol{0}$)
- Identity matrix $\boldsymbol{I}$ is ones on diagonal; zeros elsewhere. $\boldsymbol{Ix} = \boldsymbol{x}$ and $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$ for all $\boldsymbol{x}, \boldsymbol{A}$

Random Vectors

- Expectation is coordinate-wise: $\mathbb{E}[\boldsymbol{X}] = (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_n])$
- Linear transforms: $\mathbb{E}[\boldsymbol{a} + \alpha\boldsymbol{X} + \beta\boldsymbol{Y}] = \boldsymbol{a} + \alpha\mathbb{E}[\boldsymbol{X}] + \beta\mathbb{E}[\boldsymbol{Y}]$ and $\mathbb{E}[\langle \boldsymbol{a}, \boldsymbol{X}\rangle] = \langle \boldsymbol{a}, \mathbb{E}[\boldsymbol{X}]\rangle$ Does not assume independence
- PDFs work via *multiple* integrals: $\mathbb{P}(\boldsymbol{X} \in A) = \iiint_A f_{\boldsymbol{X}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. CDFs are difficult
- If joint PDF factorizes $f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$ then $X \perp\!\!\!\perp Y$ (independence)
- Marginal PDF: $f_X(x) = \int_{-\infty,\infty} f_{(X,Y)}(x,y)\,\mathrm{d}y$
- Conditional PDF: $f_{X|Y=y}(x) = f_{(X,Y)}(x,y)/f_X(x)$. General form: $f_{X|Y\in A}(x) = \int_A f_{(X,Y)}(x,y)\mathrm{d}y/\mathbb{P}(Y \in A)$

Covariance

- Covariance of two scalars: $\mathbb{C}[X,Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ (can be positive or negative)
- Self-covariance is variance: $\mathbb{C}[X,X] = \mathbb{V}[X]$
- Linear transforms: $\mathbb{C}[aX+b, cY+d] = ac\,\mathbb{C}[X,Y]$. For random vector $\boldsymbol{X}$ and fixed matrix $\boldsymbol{A}$: $\mathbb{V}[\boldsymbol{\mu} + \boldsymbol{AX}] = \boldsymbol{A}\mathbb{V}[\boldsymbol{X}]\boldsymbol{A}^T$.
- Correlation: $\rho_{X,Y} = \mathbb{C}[X,Y]/\sqrt{\mathbb{V}[X]\mathbb{V}[Y]}$
- Variance of a random vector is a *(co)variance matrix*: $\mathbb{V}[\boldsymbol{X}]_{ij} = \mathbb{C}[X_i, X_j]$
- Covariance quadratic forms give variance of linear combinations: $\mathbb{V}[\langle \boldsymbol{a}, \boldsymbol{X}\rangle] = \langle \boldsymbol{a}, \mathbb{V}[\boldsymbol{X}]\boldsymbol{a}\rangle = \sum_{ij} a_ia_j\mathbb{C}[X_i, X_j] \geq 0$
- Independence implies uncorrelated, but not the other way: $X \perp\!\!\!\perp Y \implies \mathbb{C}[X,Y] = 0 \Leftrightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

Normal Distribution

- Standard normal distribution. $Z \sim \mathcal{N}(0,1)$. Mean Zero + Variance 1
- Standard normal PDF - $\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$. Standard normal CDF $\Phi(z) = \int_{-\infty}^{z} \phi(x)\,\mathrm{d}x$ - no closed form.
- General normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ - generated by scale+shift of standard normal $X \stackrel{d}{=} \mu + \sigma Z$.
- Normal PDF via standardization (z-score): $f_X(x) = \phi(\frac{x-\mu}{\sigma}) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$. CDF: $\Phi(\frac{x-\mu}{\sigma})$.
- Multivariate normal parameterized by mean vector and (co)variance matrix: $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- Standard multi-normal: $\boldsymbol{Z} \sim \mathcal{N}_n(\boldsymbol{0}_n, \boldsymbol{I}_n)$. PDF $f_{\boldsymbol{Z}}(\boldsymbol{z}) = (2\pi)^{-n/2}e^{-\|\boldsymbol{z}\|^2/2}$.
- General multi-normal $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}$ where $\boldsymbol{\Sigma}^{1/2}$ is a matrix square root (Cholesky or symmetric).
- Bivariate normal PDF $f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2[1-\rho^2]}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right)$
- Multivariate normal: any linear combination (weighted sum) of X_i is normal.
- If $\mathbb{C}[X_i, X_j] = 0$, then $X_i \perp\!\!\!\perp X_j$ (for multi-normal, uncorrelated implies independent)
- If $\boldsymbol{Z}$ is a standard normal n-vector, $\|\boldsymbol{Z}\|^2 = \sum_{i=1}^n Z_i^2$ has a χ^2 distribution with n degrees of freedom