Math 423/533
Required Background

Probability and Statistics: key topics

- Probability (MATH 323/356)
  - Normal (Gaussian) distribution;
  - Student-t distribution;
  - Chi-squared distribution;
  - F-distribution.
- Statistics (MATH 324/357)
  - Principles of hypothesis testing: null and alternative hypothesis, test statistic, null distribution, rejection region, p-value;
  - Confidence intervals;
  - Simple linear (i.e. straight-line) regression – understanding the true relationship between predictor variable x and outcome variable y of the form

    $$y = mx + c$$

    from sample data.

General Notation for Regression Problems

- sample size n, individuals indexed by i;
- number of **predictor** (or **input**) variables, p;
- y_i: measured **outcome** for individual i;
- $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})$ – $(1 \times p)$ **row vector;**
- $\mathbf{X}$ - $(n \times p)$ **design matrix** containing all predictors for all individuals $i = 1, \ldots, n$;
  - the **rows** of $\mathbf{X}$ are denoted $\mathbf{x}_1, \ldots, \mathbf{x}_n$;
  - the **columns** of $\mathbf{X}$ are denoted $\underset{\sim}{\mathbf{x}}_1, \ldots, \underset{\sim}{\mathbf{x}}_p$;
- $\mathbf{y} = (y_1, \ldots, y_n)^\top$ – $(n \times 1)$ **column vector;**
- Y_i and $\mathbf{Y}$: random variables corresponding to outcomes.

1. **Inner product:** if $\mathbf{v}_1$ and $\mathbf{v}_2$ are $L \times 1$ vectors

$$\mathbf{v}_1 = (v_{11}, \ldots, v_{1L})^\top \qquad \mathbf{v}_2 = (v_{21}, \ldots, v_{2L})^\top$$

then

$$\mathbf{v}_1^\top \mathbf{v}_2 = v_{11}v_{21} + v_{12}v_{22} + \cdots + v_{1L}v_{2L} = \sum_{l=1}^{L} v_{1l}v_{2l}$$

Note that

$$\mathbf{v}_1^\top \mathbf{v}_2 = \mathbf{v}_2^\top \mathbf{v}_1.$$

Special case:

$$\mathbf{v}_1^\top \mathbf{v}_1 = \sum_{l=1}^{L} v_{1l}^2$$

2. **Linear independence:** the $L \times 1$ vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d$ are linearly independent if, for real-valued constants

$$\lambda_1, \lambda_2, \ldots, \lambda_d$$

we have that

$$\sum_{j=1}^{d} \lambda_j \mathbf{v}_j = \mathtt{0}_L \qquad \iff \qquad \lambda_1 = \lambda_2 = \cdots = \lambda_d = \mathtt{0}.$$

where $\mathtt{0}_L$ is the $(L \times 1)$ vector of zeros.

That is, none of the vectors can be expressed precisely as a linear combination of the others.

Linear Algebra concepts (cont.)

3. **Orthogonality:** two $L \times 1$ vectors $\mathbf{v}_1, \mathbf{v}_2$ are orthogonal if

$$\mathbf{v}_1^\top \mathbf{v}_2 = 0.$$

Matrices: Two compatible matrices $\mathbf{U}$ and $\mathbf{V}$

$\mathbf{U}$ – an $(r \times L)$ matrix $\qquad$ $\mathbf{V}$ – an $(L \times d)$ matrix

are orthogonal if

$$\mathbf{UV} = 0_{r,d}$$

where $0_{r,d}$ is the $(r \times d)$ matrix of zeros – that is, the rows of $\mathbf{U}$ are orthogonal to the columns of $\mathbf{V}$.

4. **Symmetry:** $L \times L$ matrix $\mathbf{A}$ is symmetric if

$$A_{ij} = A_{ji} \qquad \text{for all } i \text{ and } j$$

that is, the (i, j)th element is identical to the (j, i) element, for all $i, j = 1, \ldots, L$.

5. **Diagonal matrices:** $L \times L$ matrix $\mathbf{A}$ is a diagonal matrix if

$$A_{ij} = 0 \qquad \text{for all } i \neq j$$

that is, all elements off the main diagonal are zero.

Linear Algebra concepts (cont.)

6. **Positive definiteness:** An $L \times L$ matrix $\mathbf{A}$ is positive definite if for all $L \times 1$ vectors $\mathbf{v}$

$$\mathbf{v}^\top \mathbf{A}\mathbf{v} > 0.$$

- $\mathbf{A}$ is positive definite if and only if $\mathbf{A}$ is non-singular, that is, if and only if

$$\det \mathbf{A} \neq 0.$$

- if $\mathbf{A}$ is non-singular, the inverse of $\mathbf{A}$ is denoted $\mathbf{A}^{-1}$, and

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_L$$

where $\mathbf{I}_L$ is the $(L \times L)$ identity matrix.

7. **Decompositions:** If $\mathbf{A}$ is an $(L \times L)$ positive definite and symmetric matrix, then we may form the following decompositions:
   (i) Cholesky/LU decomposition:
   $$\mathbf{A} = \mathbf{L}\mathbf{L}^\top$$
   where $\mathbf{L}$ is and $L \times L$ lower-triangular matrix.
   (ii) QR decomposition:
   $$\mathbf{A} = \mathbf{Q}\mathbf{R}$$
   where $\mathbf{Q}$ is an $L \times L$ orthogonal matrix,
   $$\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_L$$
   and $\mathbf{R}$ is an $L \times L$ upper-triangular matrix.

(iii) Singular value-decomposition/eigen decomposition:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^\top$$

where $\mathbf{U}$ is an $L \times L$ orthogonal matrix,

$$\mathbf{U}^\top \mathbf{U} = \mathbf{I}_L$$

and $\mathbf{D}$ is a diagonal matrix.

Note: the QR decomposition and the SV decomposition can be applied more generally than the special case of a positive definite, symmetric matrix.

8. **Linear systems of equations:** the system of L linear equations in the quantities $v_1, \ldots, v_L$ encapsulated in the matrix form

$$\mathbf{A}\mathbf{v} = \mathbf{b}$$

for $L \times L$ matrix $\mathbf{A}$ and $L \times 1$ vector $\mathbf{b}$ has unique solution

$$\mathbf{v} = \mathbf{A}^{-1}\mathbf{b}$$

provided $\mathbf{A}$ is non-singular.

Textbook

See Appendix C of

- *Introduction to Linear Regression Analysis (5th Edition)*, by Montgomery, Peck and Vining.