- A metric $d : X \times X \to \mathbb{R}$ is a function that is used to measure distance between elements in a set X.; Defintion 2.1, pg 72
- For a metric $d : X \times X \to \mathbb{R}$, and for all $x,\, y \in X$, $d(x,\, y) = d(y,\, x)$; M1, pg 72
- For a metric $d : X \times X \to \mathbb{R}$, and for all $x,\, y \in X$, $d(x,\, y) \geq 0$; M2, pg 72
- For a metric $d : X \times X \to \mathbb{R}$, and for all $x,\, y \in X$, $d(x,\, y) = 0$ if and only if $x = y$; M3, pg 72
- For a metric $d : X \times X \to \mathbb{R}$, and for all $x,\, y,\, z \in X$, $d(x,\, z) \leq d(x,\, y) + d(y,\, z)$
  This is also known as the triangle inequality; M4, pg 72
- Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$. The l_p metric is ¡br¿ $d_p(\mathbf{x},\, \mathbf{y}) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}$; Example 2.1.2, part 3, pg 73
- Let $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$. The l_∞ metric is ¡br¿ $d_\infty(\mathbf{x},\mathbf{y}) = \max_{i=1,\,2,\ldots,\,n} |x_i - y_i|$; Example 2.1.2, part 4, pg 73
- A metric space (X, d) is a set X together with a metric d.; Definition 2.2, pg 74
- For a set $S \subset \mathbb{R}$, the least upper bound (LUB) is the smallest number z such that $z \geq x$ for every $x \in S$. The LUB of a set S is called the supremum and is denoted sup; Box 2.1, pg 74
- For a set $S \subset \mathbb{R}$, the greatest lower bound (GLB) is the largest number z such that $z \leq x$ for every $x \in S$. The GLB of a set S is called the infinum and is denoted inf; Box 2.1, pg 74
- If for every $\delta > 0$, there is an n_0 such that $d(x_n, x^*) < \delta$ for every $n > n_0$ for some fixed value x^*, then the sequence $\{x_n\}$ is said to converge to x^*.; Definition 2.10, pg 79
- If the sequence $\{x_n\}$ converges to x^*, we write $x_n \to x^*$, and we say that x^* is the limit of x_n.; Definition 2.10, pg 79
- If the sequence $\{x_n\}$ returns infinitely often to a neighborhood of a point x^*, we say that x^* is a limit point.; Definition 2.10, pg 79
- A sequence $\{x_n\}$ in a metric space (X, d) is said to be a Cauchy sequence if, for any $\epsilon > 0$, there is an $N > 0$ (which may depend on ϵ) such that $d(x_n, x_m) < \epsilon$ for every $m,\, n > N$.; Definition 2.12, pg 80
- A metric space (X, d) is complete if every Cauchy sequence in X is convergent in X.; Definition 2.13, pg 81
- A linear vector space S over a set of scalars R is a collection of objects known as vectors, together with an additive operation and a scalar multiplication operation.; Definition 2.14, pg 85
- For a linear vector space S over a set of scalars R, S forms a group under addition.; VS1, pg 85
- For a linear vector space S over a set of scalars R, for any $\mathbf{x},\, \mathbf{y} \in S$, $\mathbf{x} + \mathbf{y} \in S$.
  (The addition operation is closed.); VS1 (a), pg 85
- For a linear vector space S over a set of scalars R, there is an additive identity element in S, which is denoted as $\mathbf{0}$, such that for any $\mathbf{x} \in S$, ¡br¿$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$; VS1 (b), pg 85
- For a linear vector space S over a set of scalars R, for every element $\mathbf{x} \in S$, there is another element $\mathbf{y} \in S$ such that $\mathbf{x} + \mathbf{y} = \mathbf{0}$. The element $\mathbf{y}$ is the additive inverse of $\mathbf{x}$ and is usually denoted as $-\mathbf{x}$; VS1 (c), pg 85
- For a linear vector space S over a set of scalars R, the addition operation is associative, that is for any $\mathbf{x},\, \mathbf{y},\, \mathbf{z} \in S$,
  $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$; VS1 (d), pg 85
- For a linear vector space S over a set of scalars R, for any $a \in R$, and for any $\mathbf{x} \in S$,
  $a\mathbf{x} \in S$; VS2, pg 85
- For a linear vector space S over a set of scalars R, for any $a,\, b \in R$, and for any $\mathbf{x},\, \mathbf{y} \in S$,
  $a(b\mathbf{x}) = (ab)\mathbf{x}$; VS2, pg 85
- For a linear vector space S over a set of scalars R, for any $a,\, b \in R$, and for any $\mathbf{x},\, \mathbf{y} \in S$,
  $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$; VS2, pg 85
- For a linear vector space S over a set of scalars R, for any $a,\, b \in R$, and for any $\mathbf{x},\, \mathbf{y} \in S$,
  $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$; VS2, pg 85
- For a linear vector space S over a set of scalars R, and for any $\mathbf{x}$, there is a multiplicative identity element $1 \in R$ such that $1\mathbf{x} = \mathbf{x}$; VS3, pg 86
- For a linear vector space S over a set of scalars R, and for any $\mathbf{x}$, there is an element $0 \in R$ such that $0\mathbf{x} = 0$; VS3, pg 86
- Let S be a vector space. If $V \subset S$ such that V is a vector space, then V is said to be a subspace of S.; Definition 2.15, pg 86
- Let S be a vector space over R and let $T \subset S$. A point $\mathbf{x} \in S$ is said to be a linear combination of points in T if there is a finite set of points $\mathbf{p}_1,\, \mathbf{p}_2, \ldots,\, \mathbf{p}_m$ in T and a finite set of scalars $c_1,\, c_2, \ldots, c_m$ in R such that
  $\mathbf{x} = c_1\mathbf{p}_1,\, c_2\mathbf{p}_2, \ldots,\, c_m\mathbf{p}_m$.; Definition 2.16, pg 88
- Let S be a vector space and let $T \subset S$. The set T is linearly independent if for each finite nonempty subset of T (say $\{\mathbf{p}_1,\, \mathbf{p}_2, \ldots,\, \mathbf{p}_m\}$) the only set of scalars satisfying the equation
  $c_1\mathbf{p}_1,\, c_2\mathbf{p}_2, \ldots,\, c_m\mathbf{p}_m = 0$
  is the trivial solution $c_1 = c_2 = \cdots = c_m = 0$.; Definition 2.17, pg 88

- Let S be a vector space and let $T \subset S$. The set $\{\mathbf{p}_1,\ \mathbf{p}_2, \dots,\ \mathbf{p}_m\} \in T$ is linearly dependent if there exists a set of scalars, not all zero, such that
$c_1\mathbf{p}_1,\ c_2\mathbf{p}_2, \dots,\ c_m\mathbf{p}_m = 0$; Definition 2.17, pg 88

- Let T be a set of vectors in a vector space S. The set of vectors V that can be reached by all possible linear combinations of vectors in T is the span of the vectors.; Definition 2.18, pg 89

- Let T be a set of vectors in a vector space S and let $V \subset S$ be a subspace. If every vector $\mathbf{x} \in V$ can be written as a linear combination of vectors in T, then T is a spanning set of V; Definition 2.19, pg 90

- Let T be a set of vectors in a vector space S such that $\text{span}(T) = S$. If T is linearly independent, then T is said to be a Hamel basis for S.; Definition 2.20, pg 90

- Let S be a vector space with elements $\mathbf{x}$. A real-valued function $\|\mathbf{x}\|$ is said to be a norm if $\|\mathbf{x}\|$ satisfies the following properties:
$\|\mathbf{x}\| \geq 0$ for any $\mathbf{x} \in S$.
$\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
$\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$
$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$; Definition 2.22, pg 94

- The l_1 norm: $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$.; Example 2.3.1, pg 94

- The l_p norm: $\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$.; Example 2.3.1, pg 94

- The l_∞ norm: $\|\mathbf{x}\|_\infty = \max_{i=1,\dots,n} |x_i|$.; Example 2.3.1, pg 94

- The L_1 norm: $\|x(t)\|_1 = \int_a^b |x(t)|dt$.; Example 2.3.2, pg 95

- The L_p norm: $\|x(t)\|_p = \left(\int_a^b |x(t)|^p dt\right)^{1/p}$.; Example 2.3.2, pg 95

- The L_∞ norm: $\|x(t)\|_\infty = \sup_{t\in[a,b]} |x(t)|$.; Example 2.3.2, pg 95

- A normed linear space is a pair $(S,\ \|\cdot\|)$, where S is a vector space and $\|\cdot\|$ is a norm defined on S.; Definition 2.23, pg 95

- Let S be a vector space defined over a scalar field R. An inner product is a function $\langle\cdot,\cdot\rangle : S \times S \to R$ with the following properties: ¡ol¿ ¡li¿ $\langle\mathbf{x},\mathbf{y}\rangle = \overline{\langle\mathbf{y},\mathbf{x}\rangle}$ ¡/li¿ ¡li¿ $\langle a\mathbf{x},\mathbf{y}\rangle = a\langle\mathbf{x},\mathbf{y}\rangle$ ¡/li¿ ¡li¿ $\langle\mathbf{x}+\mathbf{y},\mathbf{z}\rangle = \langle\mathbf{x},\mathbf{z}\rangle + \langle\mathbf{y},\mathbf{z}\rangle$ ¡/li¿ ¡li¿ $\langle\mathbf{x},\mathbf{x}\rangle > 0$ if $\mathbf{x} \neq 0$, and $\langle\mathbf{x},\mathbf{x}\rangle = 0$ if and only if $\mathbf{x} = 0$ ¡/ol¿; Definition 2.26, pg 97

- For finite-dimensional vectors $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ the Euclidean inner product is $\langle\mathbf{x},\mathbf{y}\rangle = \mathbf{y}^H\mathbf{x}$.; Example 2.4.1, pg 98

- For functions defined over $\mathbb{R}$, an inner product is $\langle x(t), y(t)\rangle = \int_{-\infty}^{\infty} x(t)y(t)dt$; Example 2.4.2, pg 98

- We can use the inner product to produce a special norm, called the induced norm. Given an inner product $\langle\cdot,\cdot\rangle$ in a vector space S, we have the induced norm $\|\mathbf{x}\| = \langle\mathbf{x},\mathbf{x}\rangle^{1/2}$; Section 2.5, pg 99

- The l_p and L_p norms are only induced norms when $p = 2$.; Section 2.5, pg 99

- (Cauchy-Schwarz inequality) In an inner product space S with induced norm $\|\cdot\|$,
$|\langle\mathbf{x},\mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$ for any $\mathbf{x}, \mathbf{y} \in S$, with equality if, and only if $\mathbf{y} = a\mathbf{x}$ for some a.; Theorem 2.4, pg 100

- Vectors $\mathbf{x}$ and $\mathbf{y}$ in an inner product space are said to be orthogonal if $\langle\mathbf{x},\mathbf{y}\rangle = 0$.; Definition 2.29, pg 102

- A complete normed vector space is called a Banach space.; Definition 2.31, pg 106

- A complete normed vector space with an inner product (in which the norm is the induced norm) is called a Hilbert space.; Definition 2.31, pg 106

- A vector space equipped with an inner product is called an inner product space.; Definition 2.27, pg 97

- A matrix A is said to be positive definite (PD) if $\mathbf{x}^H A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$; Definition 3.1, pg 134

- A matrix A is said to be positive semidefinite (PSD) if $\mathbf{x}^H A\mathbf{x} \geq 0$ for all $\mathbf{x} \neq 0$; Definition 3.1, pg 134

- All diagonal elements of a positive definite (or PSD) matrix are nonnegative; Definition 3.1, pg 134

- A Hermitian matrix A is PD (or PSD) if and only if all of the eigenvalues are nonnegative. Hence, a PD matrix has a positive determinant. Hence, a PD matrix is invertable; Definition 3.1, pg 134

- A Hermitian matrix A is PD if and only if all principal minors are positive.; Definition 3.1, pg 134

- If A is PD, then the pivots obtained in the LU factorization are positive; Definition 3.1, pg 134

- If $A > 0$ and $B \geq 0$ (ie A, B are PD), then $A + B > 0$.; Definition 3.1, pg 134

- A Hermitian PD matrix A can be factored as $A = B^H B$, where B is full rank. This is a matrix square root; Definition 3.1, pg 134

- A Grammian matrix R is always positive-semidefinite. It is positive-definite if and only if the vectors $\mathbf{p}_1, \dots, \mathbf{p}_m$ are linearly independent.; Theorem 3.1, pg 134

- Let $\mathbf{p}_1, \dots, \mathbf{p}_m$ be data vectors in a vector space S. Let $\mathbf{x} \in S$. In the representation
$$\mathbf{x} = \sum_{i=1}^{m} c_i\mathbf{p}_i + \mathbf{e} = \hat{\mathbf{x}} + \mathbf{e},$$
the induced norm of the error vector $\|\mathbf{e}\|$ is minimized when the error $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ is orthogonal to each of the data vectors. i.e.
$$\left\langle \mathbf{x} - \sum_{i=1}^{m} c_i\mathbf{p}_i, \mathbf{p}_j \right\rangle = 0$$
for $j = 1, 2, \dots, m$.; Theorem 3.2, pg 135

- The optimal (least-squares) coefficients $\mathbf{c}$ are

$$\mathbf{c} = (A^H A)^{-1} A^H \mathbf{x}$$

.; Equation 3.19, pg 139

- A transformation $A : X \to Y$, where X and Y are vector spaces over a ring R is said to be linear if for every $x_1, x_2 \in X$ and all scalars $\alpha_1, \alpha_2 \in R$ $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)$; Defintion 4.1, pg 230

- A functional $f\colon\ X \to \mathbb{R}$ is a mapping from a vector space to a real scalar value.; Definition 4.2, pg 231

- An operator norm provides an indication of the maximal amount of change of length of a vector that it operates on.; Section 4.2, pg 232

- The p operator norm of $A : X \to Y$ is $\|A\|_p = \sup_{x\in X, \neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{x\in X, \|x\|=1} \|Ax\|_p$; Section 4.2, pg 232

- The operator norm $\|A\|_{op}$ can also be defined with the inf: $\|A\|_{op} = \inf\{c \geq 0 : \|Ax\| \leq c\|x\| \text{ for all } x \in X\}$; Wikipedia, Equivalent Definitions of Operator Norms

- If the norm of a transformation is finite, the transformation is said to be bounded.; Defintion 4.3, pg 233

- A linear operator $A : X \to Y$ is bounded if and only if it is continuous.; Theorem 4.1, pg 233

- Let $A : X \to Y$ be a linear operator. If X is finite dimensional, then A is continuous.; Theorem 4.2, pg 233

- For a scalar x where $|x| < 1$, $1+x+x^2+\cdots = \sum_{i=0}^{\infty} x^i = \frac{1}{1-x} = (1-x)^{-1}$; Neumann expansion, pg 235

- Suppose $\|\cdot\|$ is a norm satisfying the submultiplicative property and A is an operator with $\|A\| < 1$. Then, $(I - A)^{-1} = \sum_{i=0}^{\infty} A^i$; Theorem 4.3, pg 235

- The p norms satisfy the submultiplicative property; pg 233

- The submultiplicative property $\|AB\| \leq \|A\|\|B\|$; pg 233

- For a matrix A $\|A\|_\infty = \max_i \sum_j |a_{ij}|$; Equation 4.5, pg 235

- For a matrix A $\|A\|_1 = \max_j \sum_i |a_{ij}|$; Equation 4.6, pg 236

- For a matrix A, $\|A\|_\infty$ is the largest row sum; Equation 4.5, pg 235

- For a matrix A, $\|A\|_1$ is the largest column sum; Equation 4.6, pg 236

- The Frobenius norm (sum form) $\|A\|_F = \left(\sum_i \sum_j |a_{ij}|^2\right)^{1/2}$; pg 237

- The Frobenius norm (trace form) $\|A\|_F = \mathrm{tr}(A^H A)^{1/2}$; pg 237

- The adjoint is defined for $A : X \to Y$, a bounded linear operator where X, Y are Hilbert spaces.; Definition 4.4, pg 237

- The adjoint of the operator $A : X \to Y$ is the operator $A^* : Y \to X$; Defintion 4.4, pg 237

- The adjoint of the operator $A : X \to Y$ is $\langle Ax, y\rangle = \langle x, A^* y\rangle$ for all x, y; Defintion 4.4, pg 237

- An operator A is self-adjoint if $A^* = A$; Defintion 4.4, pg 237

- The adjoint of a matrix is the conjugate transpose of the matrix.; pg 238

- A real matrix which is self-adjoint is said to be symmetric; pg 238

- A complex matrix which is self-adjoint is said to be Hermitian; pg 238

- $(A_2 A_1)^* = A_1^* A_2^*$; Property 3, pg 238

- $(A_1 + A_2)^* = A_1^* + A_2^*$; Property 1, pg 238

- $(\alpha A)^* = \bar{\alpha} A^*$; Property 2, pg 238

- If A has an inverse, then $(A^{-1})^* = (A^*)^{-1}$; Property 4, pg 238

- The space spanned by the columns of a matrix is called the column space or range of the matrix.; Definition 4.5, pg 241

- The range of a matrix A is denoted $\mathcal{R}(A)$; Defintion 4.5, pg 241

- The equation $A\mathbf{x} = \mathbf{b}$ has a solution only if $\mathbf{b}$ lies in the column space of A; Box, pg 241

- The nullspace of a linear operator $A : X \to Y$ consists of all vectors $x \in X$ such that $Ax = 0$; Defintion 4.6, pg 242

- The nullspace of A is denoted as $\mathcal{N}(A)$; Defintion 4.6, pg 242

- The dimension of $\mathcal{N}(A)$ is called the nullity of A; Defintion 4.6, pg 242

- For a linear operator $A : X \to Y$, the range of the adjoint is denoted $\mathcal{R}(A^*)$; pg 242

- For a linear operator $A : X \to Y$, the nullspace of the adjoint is denoted $\mathcal{N}(A^*)$ and is also called the left nullspace; pg 242

- For a linear operator $A : X \to Y$, $\mathcal{R}(A) \subset Y$; Equation 4.19 pg 242

- For a linear operator $A : X \to Y$, $\mathcal{N}(A) \subset X$; Equation 4.19 pg 242

- For a linear operator $A : X \to Y$, $\mathcal{R}(A^*) \subset X$; Equation 4.19 pg 242

- For a linear operator $A : X \to Y, \mathcal{N}(A^*) \subset Y$; Equation 4.19 pg 242
- Let $A : X \to Y$ be a bounded linear operator with X, Y Hilbert spaces, and let $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ be closed. Then $[\mathcal{R}(A)]^\perp = \mathcal{N}(A^*)$; Equation 4.20, pg 242
- Let $A : X \to Y$ be a bounded linear operator with X, Y Hilbert spaces, and let $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ be closed. Then $[\mathcal{R}(A^*)]^\perp = \mathcal{N}(A)$; Equation 4.21, pg 243
- Let $A : X \to Y$ be a bounded linear operator with X, Y Hilbert spaces, and let $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ be closed. Then $\mathcal{R}(A) = [\mathcal{N}(A^*)]^\perp$; Equation 4.20, pg 242
- Let $A : X \to Y$ be a bounded linear operator with X, Y Hilbert spaces, and let $\mathcal{R}(A)$ and $\mathcal{R}(A^*)$ be closed. Then $\mathcal{R}(A^*) = [\mathcal{N}(A)]^\perp$; Equation 4.21, pg 243
- A matrix A is said to have a left inverse if there is a matrix B such that $BA = I$; Definition 4.8, pg 247
- A matrix A is said to have a right inverse if there is a matrix B such that $AB = I$; Definition 4.8, pg 247
- $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(A^*))$; Notes from 17-Oct
- A $n \times n$ matrix is invertible if $\mathcal{N}(A) = \{0\}$; Test 1, pg 248
- A $n \times n$ matrix is invertible if $\text{rank}(A) = n$; Test 2, pg 248
- A $n \times n$ matrix is invertible if the rows and columns of A are linearly independent; Test 3, pg 248
- A $n \times n$ matrix is invertible if the determinant of A is nonzero; Test 4, pg 248
- A $n \times n$ matrix is invertible if there are no zero eigenvalues of A; Test 5, pg 248
- A $n \times n$ matrix is invertible if $A^H A$ is positive definite; Test 6, pg 248
- A matrix A is nonsingular if $A\mathbf{x} = \mathbf{0}$ has only the solution $\mathbf{x} = \mathbf{0}$; Defintion 4.9, pg 248
- The condition number of a matrix A is $\kappa(A) = \|A\|\|A^{-1}\|$; pg 254
- Rule of thumb with condition number¡br¿Let $p = \log_{10}(\kappa(A))$.
  If the solution is computed to n decimal places, then only about $n - p$ places can be considered to be accurate.; pg 256
- The LU factorization.
  $PA = LU$; Equation 5.2, pg 276
- In the LU factorization, $PA = LU$, the matrix A must be square;Equation 5.2, pg 276
- In the LU factorization, $PA = LU$, L is a lower-triangular matrix with ones on the main diagonal; Equation 5.2, pg 276
- In the LU factorization, $PA = LU$, U is an upper-triangular matrix; Equation 5.2, pg 276
- In the LU factorization, $PA = LU$, P is a permutation matrix; Equation 5.2, pg 276
- The Cholesky factorization.
  $A = LL^H$; pg 283
- The Cholesky factorization can be interpreted as a Matrix square-root; pg 283
- In the Cholesky factorization, $A = LL^H$, L is lower-triangular; pg 283
- In the Cholesky factorization, $A = LL^H$, A must be Hermitian, square, and positive-definite; pg 283
- For a unitary (or orthogonal) matrix Q,
  $Q^H Q = QQ^H = I$.; Definition 5.1, pg 285
- A matrix Q where $Q^H Q = I$ is called unitary if its elements are complex and orthogonal if its elements are real.; Definition 5.1, pg 285
- For $\mathbf{y} = Q\mathbf{x}$, $\|\mathbf{y}\| = \|\mathbf{x}\|$ if and only if Q is unitary; Lemma 5.1, pg 385
- In the QR factorization,
  $A = QR$
  where A is an arbitrary dimension; pg 286
- In the QR factorization, $A = QR$ with A an $m \times n$ matrix,
  Q is orthogonal and $m \times m$::dimension; pg 286
- In the QR factorization, $A = QR$ with A an $m \times n$ matrix,
  R is upper triangular and $m \times n$::dimension; pg 286
- An eigenvalue and an eigenvector of a matrix A is a scalar λ and a vector $\mathbf{x}$ that satisfy
  $A\mathbf{x} = \lambda\mathbf{x}$; Equation 6.5, pg 306
- The eigenvectors of A are those vectors that are scaled and not changed in direction.; pg 306
- The characteristic polynomial of A is
  $\det(\lambda I - A)$. ¡br¿The roots of the characteristic polynomial are the eigenvalues of A; Definition 6.2, pg 306
- The set of roots of the characteristic equation is called the spectrum of A and is denoted $\lambda(A)$; Definition 6.2, pg 306
- If the eigenvalues of an $m \times m$ matrix A are all distinct, then the eigenvectors are linearly independent; Lemma 6.1, pg 308
- If the eigenvectors of the matrix A are linearly independent, then A can be diagonalized as
  $A = S\Lambda S^{-1}$ where S is a matrix whose columns are the eigenvectors of A and Λ is a diagonal matrix with the eigenvalues of A on the diagonal.; Equation 6.11, pg 309
- Every self-adjoint matrix A can be diagonalized by a unitary (orthogonal) matrix U:
  $A = U\Lambda U^H$; Theorem 6.2, pg 313

- The singular value decomposition (SVD).
  Every matrix $A \in \mathbb{C}^{m \times n}$ can be factored as
  $A = U\Sigma V^H$; Theorem 7.1, pg 369
- In the singular value decomposition, for a $m \times n$ matrix $A = U\Sigma V^H$
  U is $m \times m$::dimension and $U^H U = I$.; Theorem 7.1, pg 369
- In the singular value decomposition, for a $m \times n$ matrix $A = U\Sigma V^H$
  V is $n \times n$::dimension and $V^H V = I$.; Theorem 7.1, pg 369
- In the singular value decomposition, for a $m \times n$ matrix $A = U\Sigma V^H$
  Σ is $m \times n$::dimension and diagonal; Theorem 7.1, pg 369
- In the singular value decomposition, for a $m \times n$ matrix $A = U\Sigma V^H$
  the singular values σ are the eigenvalues of $A^H A$ and AA^H; Theorem 7.1, pg 369
- The rank of a matrix is the number of nonzero singular values.; pg 372
- The singular value decomposition of a matrix A can be written as $A = U\Sigma V^H = [U_1 U_2] \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix}$ where Σ_1 is square::shape and has the singular values of A on the diagonal and where Σ_2 is entirely zeros.; pg 371
- Fundamental subspaces and the SVD
  $\mathcal{R}(A) = \text{span}(U_1)$; Equation 7.7, pg 372
- Fundamental subspaces and the SVD
  $\mathcal{R}(A^H) = \text{span}(V_1)$; Equation 7.7, pg 372
- Fundamental subspaces and the SVD
  $\mathcal{N}(A) = \text{span}(V_2)$; Equation 7.7, pg 372
- Fundamental subspaces and the SVD
  $\mathcal{N}(A^H) = \text{span}(U_2)$; Equation 7.7, pg 372
- The pseudoinverse of A can be written using the SVD as $A^\dagger = V\Sigma^\dagger U^H$; Equation 7.11, pg 374

IID Independent and identically distributed

Grammian matrix

$$R = \begin{bmatrix} \langle \mathbf{p}_1, \mathbf{p}_1 \rangle & \cdots & \langle \mathbf{p}_m, \mathbf{p}_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{p}_1, \mathbf{p}_m \rangle & \cdots & \langle \mathbf{p}_m, \mathbf{p}_m \rangle \end{bmatrix}$$

Where $\mathbf{p}_i$ are vectors Eq. 3.7

Projection matrix

$$P_A = A(A^H A)^{-1} A^H$$

The matrix P_A projects onto the range of A. pg 139