BASIC LINEAR ALGEBRA NOTES

4/29/2010

Nándor Sieben

Contents

1. Systems of linear equations

1. **linear equation**: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
   variables: $x_1, \ldots, x_n$
   coefficients: $a_1, \ldots, a_n$
   main coefficient: a_1
   constant term: b
2. **linear system**: m equations, n unknowns

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

3. **solution**: n-tuple $(x_1, \ldots, x_n)$ satisfying all equations
4. **consistent system**: has a solution
5. **inconsistent system**: has no solution
6. **solution set**: set of all solutions
7. **equivalent systems**: have the same solution set
8. **elementary (row) operations on equations**: make equivalent systems
   (i) multiply an equation by a nonzero constant
   (ii) interchange two equations
   (iii) add a constant multiple of an equation to another
9. **elimination**: use elementary operations to eliminate unknowns
10. **fact**: a linear system has no solution, exactly one solution or infinitely many solutions
11. **parameters**: used to describe infinitely many solutions
12. **homogeneous system**: constant terms are 0 (consistent)
13. **trivial solution**: all variables are 0

2. Matrices of a system

1. **coefficient matrix**:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \\ a_{m1} & & a_{mn} \end{bmatrix}$$

2. **constant vector**: $b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ **unknown vector**: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$
3. **augmented matrix**:

$$[A \quad b] = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & & \\ a_{m1} & & a_{mn} & b_m \end{bmatrix}$$

3. Gauss elimination

1. **elementary row operations**: (ero) correspond to elementary operations on equations
   (i) multiply a row by a nonzero constant $r_i \leftarrow cr_i$
   (ii) interchange two rows $r_i \leftrightarrow r_j$
   (iii) add a multiple of a row to another row $r_i \leftarrow r_i + cr_j$
2. **row equivalent matrices**: one can be gotten from the other by elementary row operations
3. **fact**: linear systems with row equivalent augmented matrices have the same solution set
4. **echelon matrix**: the number of leading zeros is strictly increasing in each row until you get all 0 rows
5. **Gauss elimination**: use elementary row operations to get echelon form
6. **leading entry**: first nonzero entry in a row
7. **leading (pivot) column**: column containing a leading entry
8. **leading variable**: a variable corresponding to a leading column
9. **free variable**: not leading
10. **free column**: not leading
11. **back substitution**: get solution set from echelon form
    (i) set free variables equal to parameters
    (ii) solve last nonzero equation for leading variable
    (iii) substitute into preceding equation
    (iv) continue
12. **reduced echelon matrix**:
    (i) echelon matrix
    (ii) every leading entry is 1
    (iii) every leading entry is the only nonzero entry in it's column
13. **Gauss-Jordan elimination**: use elementary row operations to get reduced echelon form
14. **fact**: every matrix is row equivalent to a unique reduced echelon matrix
15. **fact**: system with square coefficient matrix A has unique solution iff A is row equivalent to I
16. **fact**: system with more unknowns than equations is inconsistent or has infinitely many solutions

4. Matrices

1. **matrix**: rectangular array of numbers
2. **notation**: $A = [\, a_{ij} \,]$
3. **scalar**: real number
4. **size of a matrix**: $\text{size}(A) = m \times n$ if m rows and n columns
5. **square matrix**: $m = n$
6. **diagonal matrix**: $D = [\, d_{ij} \,]$ $d_{ij} = 0$ if $i \neq j$
7. **zero matrix**: O all entries o_{ij} are 0
8. **identity matrix**: $I = [\, \delta_{ij} \,]$ $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$
9. **(column) vector**: has size $n \times 1$
10. **row vector**: has size $1 \times n$
11. **n-tuple**: $(a_1, \dots, a_n) \equiv \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \neq [\, a_1 \quad \cdots \quad a_n \,]$ slightly abusive identification
12. $\mathbf{R}^n$: set of n-tuples, $\mathbf{R}^2$ = plane, $\mathbf{R}^3$ = space
13. $\mathbf{R}^{m \times n}$: set of $m \times n$ matrices, $\mathbf{R}^{n \times 1}$ is identified with $\mathbf{R}^n$
14. **basic unit vectors**: $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ (1 in j-th position), column vectors of $I = [\, e_1 \quad \cdots \quad e_n \,]$
15. **column vectors**: $A = [\, c_1 \quad \cdots \quad c_n \,]$

5. Matrix operations

1. **matrix addition**: $A+B=[\,a_{ij}+b_{ij}\,]$ if A, B have the same size
2. **matrix subtraction**: $A-B=[\,a_{ij}-b_{ij}\,]$
3. **scalar multiplication**: $cA=[\,ca_{ij}\,]$
4. **negative matrix**: $-A=(-1)A$
5. **properties**:
   $A+B=B+A$ commutative
   $A+(B+C)=(A+B)+C$ associative
   $c(A+B)=cA+cB$ distributive
   $(c+d)A=cA+dA$ distributive
   $(cd)A=c(dA)$ associative
6. **matrix multiplication**: $C=AB$, $\text{size}(C)=m\times n$, $\text{size}(A)=m\times p$, $\text{size}(B)=p\times n$
   $c_{ij}=\sum_{k=1}^{p}a_{ik}b_{kj}=(i\text{-th row of }A)\cdot(j\text{-th column of }B)$
7. **properties**:
   $A(BC)=(AB)C$ associative
   $A(B+C)=AB+AC$ distributive
   $(A+B)C=AC+BC$ distributive
   $c(AB)=(cA)B=A(cB)$
8. **warning**:
   $AB\neq BA$ in general
   $AB=AC\not\Rightarrow B=C$
   $AB=O\not\Rightarrow A=O$ or $B=O$
9. **transpose**: $A^T=[\,b_{ij}\,]$ where $b_{ij}=a_{ji}$
10. **properties**:
    $(A^T)^T=A$
    $(A+B)^T=A^T+B^T$
    $(cA)^T=cA^T$
    $(AB)^T=B^TA^T$
11. **trace of a square matrix**: sum of the diagonal entries $\text{tr}(A)=a_{1,1}+\cdots+a_{n,n}$
12. **fact**: product of diagonal matrices is diagonal
13. **matrix form of linear system**: $Ax=b$, $A=[\,a_{ij}\,]$, $x=(x_1,\dots,x_n)$, $b=(b_1,\dots,b_n)$
14. **linear combination**: of objects v_i is a finite sum of scalar multiples of the objects $\sum_{i=1}^{n}c_iv_i$, $c_i\in\mathbf{R}$
15. **fact**: Ax is the linear combination $x_1c_1+\cdots x_nc_n$ of the columns of A
16. **span**: of objects v_i is the set of linear combinations of the objects $\text{span}\{v_1,\dots,v_n\}=\{\sum_{i=1}^{n}c_iv_i\mid c_i\in\mathbf{R}\}$
17. **fact**: solution set of homogeneous system is the span of particular solutions (one for each parameter)

6. Inverse matrix

1. A **invertible**: $\exists B$ such that $AB=BA=I$
   B is the **inverse** of A (A is also the inverse of B)
2. **properties**:
   invertible $\Rightarrow$ square
   inverse is unique if exists, notation A^{-1}
   $(A^{-1})^{-1}=A$
   $(AB)^{-1}=B^{-1}A^{-1}$
   $(A^T)^{-1}=(A^{-1})^T$
   if A is invertible then $Ax=b$ has unique solution $x=A^{-1}b$
3. **fact**: $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$ is invertible iff $ad\neq bc$, $A^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d&-b\\-c&a\end{bmatrix}$
4. **elementary matrix**: $I\overset{\text{ero}}{\mapsto}E$ single elementary row operation

5. **properties**:
   $I \overset{\text{ero}}{\mapsto} E$ implies $A \overset{\text{ero}}{\mapsto} EA$ equivalently $[\, I \quad A \,] \overset{\text{ero}}{\mapsto} [\, E \quad EA \,]$
   $I \overset{\text{iero}}{\mapsto} E^{-1}$ inverse ero
6. **fact**: A invertible iff A row equivalent to I
7. **fact**: A, B row equivalent iff $A = E_1 \cdots E_n B$, for E_i elementary
8. **algorithm for A^{-1}**: $[\, A \quad I \,] \overset{\text{ero's}}{\mapsto} [\, I \quad A^{-1} \,]$
   more generally $[\, A \quad B \,] \overset{\text{ero's}}{\mapsto} [\, I \quad A^{-1}B \,]$

7. Determinants

1. **notation**: $A = [\, a_{ij} \,]$ $n \times n$
2. 1×1 **matrix**: $\det [\, a \,] = a$
3. 2×2 **matrix**: $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$
4. **notation**: A_{ij} =submatrix after deletion of i-th row and j-th column
5. **ij-th cofactor of A**: $C_{ij} = (-1)^{i+j} \det A_{ij}$
6. **chess board rule**: $\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$ $(-1)^{i+j}$
7. **inductive definition**: $\det A = \sum_{j=1}^{n} a_{1j} C_{1j}$
   cofactor expansion along first row
8. **cofactor expansion**:
   along i-th row $\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$
   along j-th column $\det A = \sum_{i=1}^{n} a_{ij} C_{ij}$
9. **elementary row operations**: $A \overset{\text{ero}}{\mapsto} B$
   $r_i \leftarrow c r_i$: $\det B = c \cdot \det A$
   $r_i \leftrightarrow r_j$: $\det B = -\det A$
   $r_i \leftarrow r_i + c r_j$: $\det B = \det A$
10. **properties**:
   A triangular implies $\det(A) = a_{11} \cdots a_{nn}$
   $\det I = 1$
   $r_i = r_j$ implies $\det A = 0$
   $$\det \begin{bmatrix} r_1 \\ \vdots \\ r_i + r_i' \\ \vdots \\ r_n \end{bmatrix} = \det \begin{bmatrix} r_1 \\ \vdots \\ r_i \\ \vdots \\ r_n \end{bmatrix} + \det \begin{bmatrix} r_1 \\ \vdots \\ r_i' \\ \vdots \\ r_n \end{bmatrix}$$
   $\det kA = k^{\text{size}(A)} \det A$
   $\det A^T = \det A$
   $\det(AB) = \det A \cdot \det B$
   $\det(A^{-1}) = \dfrac{1}{\det(A)}$
   A invertible iff $\det A \neq 0$
11. **Cramer's rule**: $\det A \neq 0$ implies solution of $Ax = b$ is
   $x_i = \dfrac{\det A_i}{\det A}$ where A_i comes from A after replacing i-th column by b
12. **classical adjoint (adjugate) of A**: $\operatorname{adj} A = [\, C_{ij} \,]^T$ transpose of matrix of cofactors
13. **adjoint formula for inverse**: $A^{-1} = \dfrac{\operatorname{adj} A}{\det A}$

8. Vector spaces

1. **vector space**: set V of vectors with verctor addition and scalar multiplication satisfying for all $u, v, w \in U$ and $c, d \in \mathbf{R}$
   i) $u + v = v + u$
   ii) $(u + v) + w = u + (v + w)$
   iii) $\exists \underline{0} \in V$, $u + \underline{0} = u$
   iv) $\exists - u \in V$, $u + (-u) = \underline{0}$
   v) $c(u + v) = cu + cv$
   vi) $(c + d)u = cu + du$
   vii) $c(du) = (cd)u$
   viii) $1u = u$
2. **examples**: $\mathbf{R}^n$, $\mathbf{R}^{m\times n}$, $\mathbf{P}$ polynomials, $\mathbf{P}_n$ polynomials with degree less than n, sequences, sequences converging to 0, functions on $\mathbf{R}$, $C(\mathbf{R})$ continuous functions on $\mathbf{R}$, solutions of homogeneous systems
3. **subspace of** V: subset W of V that is a vector space with same operations
4. **proper subspace of** V: subspace but not $\{\underline{0}\}$ and not V
5. **examples**:
   $W = \{0\}$ and $W = V$, subspaces of V
   $W =$lines through origin, subspace of $V = \mathbf{R}^2$
   $W =$planes through origin, subspace of $V = \mathbf{R}^3$
   $W =$diagonal $n \times n$ matrices, subspace of $V = \mathbf{R}^{n\times n}$
   $W = \operatorname{span}\{v_1, \ldots, v_n\}$, subspace of V where $v_1, \ldots, v_n \in V$
   $W =$convergent sequences, subspace of $V =$sequences
   $W =$continuous functions on $\mathbf{R}$, subspace of $V =$functions on $\mathbf{R}$
6. **fact**: subset W of V is a subspace of V iff
   nonempty: $W \neq \emptyset$
   closed under addition: $\forall u, v \in W$, $u + v \in W$
   closed under scalar multiplication: $\forall c \in \mathbf{R}\ \forall u \in W$, $cu \in W$

9. Linear independence

1. $v_1, \ldots, v_n$ **linearly independent**: $\sum_{i=1}^{n} c_i v_i = \underline{0}$ implies $\forall i$, $i = 0$
2. **linearly dependent**: not independent
3. **parallel vectors**: one is scalar multiple of the other
   notation $u \| v$
4. **properties**:
   u, v linearly independent iff $u \| v$
   vectors are dependent iff one of them is linear combination of the others
   subset of lineraly independent set is linearly independent
   columns of matrix A are independent iff $AX = 0$ has only trivial solution
   columns of square matrix A are independent iff A invertible iff $\det A \neq 0$
   $v_1, \ldots, v_n$ independent, $v_{n+1} \notin \operatorname{span}\{v_1, \ldots, v_n\}$ implies $v_1, \ldots v_{n+1}$ independent
   $v_1, \ldots, v_n$ independent, $\sum_{i=1}^{n} c_i v_i = \sum_{i=1}^{n} d_i v_i$ implies $\forall i$, $c_i = d_i$
   rows of row echelon matrix are independent
   leading columns of echelon matrix are independent

10. Bases

1. S **spans** W: span$S = W$
   S is a spanning set of W
2. **basis of** V: linearly independent spanning set of V
   maximal independent set in V
   minimal spanning set of V
   spanning set containing dim (V) vectors
   independent set containing dim (V) vectors
3. **standard bases** $E = \{e_1, \ldots, e_n\}$ **for** V:
   $\{(1,0),(0,1)\}$ for $\mathbf{R}^2$
   $\{1, x, x^2\}$ for $\mathbf{P}_3(x)$
   $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ for $\mathbf{R}^{2\times 2}$
4. **replacement theorem**: span$S = V$, $T \subseteq V$, $|T| > |S|$ implies T dependent
5. **dimension of** V:
   all bases of V has same number of vectors
   dim$V$ =number of vectors in a basis of V
6. **examples**:
   dim$\mathbf{R}^n = n$
   dim$\{\underline{0}\} = 0$
   dim$\mathbf{R}^{m\times n} = mn$
   dim$\mathbf{P}_n(x) = n$
   dim$\mathbf{P}(x) = \infty$
   dim(span$\{u\}) = 1$
7. **properties**:
   W proper subspace of V implies dim$W <$ dim$V$
   independent subset of V can be extended to a basis of V
   spanning set of V contains a basis of V

11. Row, column and null spaces

1. **notation**: size$A = m \times n$
2. **row space of** A: Row$A$ =subspace of $\mathbf{R}^m$ spanned by rows of A
3. **row rank of** A: dim Row$A$
4. **column space of** A: Col$A$ =subspace of $\mathbf{R}^n$ spanned by columns of A
5. **column rank of** A: dim Col$A$
6. **algorithm for basis of** Col$A$:
   (i) reduce A to echelon form B
   (ii) take columns of A corresponding to leading columns of B
7. **algorithm for basis of** Row$A$: find basis for Col(A^T)
8. **fact**: row rank A equals column rank A
   rank A: this common value
9. **null space of** A: Null$A = \{x \mid Ax = 0\}$ =solution set of homogeneous system, subspace of $\mathbf{R}^n$
10. **properties**:
    Null$(A) =$ Row$(A)^\perp$
    Null$(A^T) =$ Col$(A)^\perp$
    A, B row equivalent implies Row$A =$ Row$B$
    A, B row equivalent implies colums of A and columns of B have the same dependence relations
    $Ax = b$ consistent iff $b \in$ Col$A$
    rank$A +$ dim Null$A = n$

12. Coordinates

1. **notation**: $B = \{b_1, \ldots, b_n\}$, $D = \{d_1, \ldots, d_n\}$ bases for V, $E = \{e_1, \ldots, e_n\}$ standard basis for V
2. **fact**: each $v \in V$ can be written uniquily as $v = c_1 b_1 + \cdots + c_n b_n$
3. **coordinates of v in basis B**: $[v]_B = (c_1, \ldots, c_n)$ if $v = \sum_{i=1}^n c_i b_i$
4. **huge fact**: $v \mapsto [v]_B : V \to \mathbf{R}^n$ is an isomorphism ($\mathbf{R}^n$ are the 'only' finite dimensional vector spaces)
5. **transition matrix from basis B to basis D**: $T_B^D = [\,[b_1]_D \quad \cdots \quad [b_n]_D\,]$ square matrix
6. **properties**:
   $[v]_D = T_B^D [v]_B$
   $T_B^D = (T_D^B)^{-1}$
   $T_B^D = T_E^D T_B^E = (T_D^E)^{-1} T_B^E$
   $[\,T_D^E \quad T_B^E\,] \overset{\text{eros}}{\mapsto} [\,I \quad T_B^D\,]$
7. **algorithm for finding a basis for $W = \operatorname{span}\{v_1, \ldots, v_n\}$ in V**:
   (i) find a bases B for V (use standard if possible)
   (ii) put the coordinates of the v_i's as columns for a matrix A
   (iii) reduce A to echelon form B
   (iv) take columns of A corresponding to leading columns of B
   (v) use these columns as coordinates to build the basis of W
8. **algorithm for extending a linearly independent set $\{v_1, \ldots, v_n\}$ to get a basis**:
   use the previous algorithm to find a basis for $\operatorname{span}\{v_1, \ldots, v_n, e_1, \ldots, e_n\}$

13. Linear transformations

1. **notation**: $B = \{b_1, \ldots, b_m\}$ basis for V, $D = \{d_1, \ldots, d_n\}$ basis for W, E standard basis for V
2. **linear transformation**: $L : V \to W$ such that for all $u, v \in V$, $\alpha \in \mathbf{R}$
   i) $L(u + v) = L(u) + L(v)$ additive
   ii) $L(\alpha u) = \alpha L(u)$ multiplicative
3. **kernel**: $\ker L = \{v \in V \mid L(v) = \underline{0}\}$
4. **image or range**: $\operatorname{im} L = \operatorname{ran} L = \{L(v) \mid v \in V\} = \operatorname{ran} L = \operatorname{span}\{Lb_1, \ldots, Lb_n\}$
5. **L is one-to-one (1-1)**: $L(u) = L(v)$ implies $u = v$
6. **L is onto W**: $\operatorname{ran} L = W$
7. **L is an isomorphism**: if L is one-to-one and onto
8. **properties**:
   $L(\underline{0}) = \underline{0}$
   $\ker L$ subspace of V
   $\operatorname{ran} L$ subspace of W
   L is 1-1 iff $\ker L = \{\underline{0}\}$
9. **matrix of L**: $[L]_B^D = [\,[Lb_1]_D \quad \cdots \quad [Lb_m]_D\,]$
10. **properties**:
    $[L]_B^D = T_E^D [L]_B^E = (T_D^E)^{-1} [L]_B^E$
    $[L]_B^D = (T_D^E)^{-1} [L]_E^E T_B^E$ if $V = W$
    $[Lv]_D = [L]_B^D [v]_B$
    $[L^{-1}]_D^B = ([L]_B^D)^{-1}$
11. **R, S are similar matrices**: $S = P^{-1} R P$ for some P (P is a transition matrix)
12. **fact**: R, S are similar iff $R = [L]_B^B$, $S = [L]_D^D$ where $V = W$
13. **rank of L**: $\operatorname{rank} L = \dim \operatorname{ran} L$
14. **properties**: $M = [L]_B^D$
    $[\operatorname{ran} L]_D = \operatorname{Col} M$
    $[\ker L]_B = \operatorname{Null} M$
    $\operatorname{rank} L = \operatorname{rank} M$
    $\dim \ker L = \dim \operatorname{Null} M$
15. **dimension theorem**: $\operatorname{rank} L + \dim \ker L = \dim V$

14. Eigenvalues and eigenvectors

1. **notation**: $L : V \to V$ linear transformation, $A = [L]_B^B$ matrix of L, $x = [u]_B$ coordinates of u
2. **eigenvalue problem**:
   transformation version $L(u) = \lambda u$, $u \neq \underline{0}$
   eigenvalue: λ
   eigenvector of L associated to λ: u
   eigenspace associated to λ: $E_\lambda = \ker(L - \lambda \mathrm{id})$
   matrix version $Ax = \lambda x$, $x \neq \underline{0}$
   eigenvalue: λ
   eigenvector of A associated to λ: x
   eigenspace associated to λ: $E_\lambda = \mathrm{Null}(A - \lambda I)$
3. **characteristic polynomial**: $\det(A - \lambda I)$
   if $A \sim B$ then $\mathrm{charpoly}(A) = \mathrm{charpoly}(B)$
4. **characteristic equation**: λ eigenvalue of A iff $\det(A - \lambda I) = 0$
5. **algebraic multiplicity of λ**: multiplicity of λ as a root of the charachteristic polynomial
6. **geometric multiplicity of λ**: $\dim E_\lambda$

15. Diagonalization

1. A **diagonalizable**: A similar to diagonal matrix D, $D = P^{-1}AP$
2. **fact**: $D = P^{-1}AP$ implies
   $P = [\, v_1 \quad \cdots \quad v_n \,]$
   $D = [\, d_{ij} \,]$, $d_{ij} = \begin{cases} \lambda_i & i = j \\ 0 & i \neq j \end{cases}$
   $Av_i = \lambda_i v_i$
   $\{v_1, \dots, v_n\}$ is a basis of eigenvectors with associated eigenvalues in the diagonal of D
3. **properties**:
   A is diagonalizable iff for each eigenvalue the algebraic and geometric multiplicities are the same
   if $v_1, \dots, v_n$ eigenvectors associated to distinct eigenvalues then they are independent
   if $\mathrm{size} A = n \times n$ and A has n distinct eigenvalues then A diagonalizable
   $\lambda_1, \dots, \lambda_n$ distinct eigenvalues, $B_1, \dots, B_n$ bases for eigenspaces implies $B_1 \cup \dots \cup B_n$ is independent
4. **algorithm for diagonalization**:
   (i) solve charachteristic equation to find eigenvalues
   (ii) for each eigenvalue λ find basis B_λ of associated eigenspace E_λ
   (iii) if the union $\cup B_\lambda$ of the bases is not a basis for the vectorspace than not diagonalizable
   (iv) build P from the eigenvectors as columns
   (v) build D from the corresponding eigenvalues

16. Inner product

1. **inner product**: a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbf{R}$ satisfying
   (i) $\langle u, v \rangle = \langle v, u \rangle$
   (ii) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$
   (iii) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
   (iv) $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ iff $u = \underline{0}$
2. **examples of inner products**:
   dot product (standard inner product) on $\mathbf{R}^n$: $\langle u, v \rangle = u \cdot v = \sum_{i=1}^n u_i v_i = u^T v = u^T I v$
   standard inner product on $C[0,1]$: (continuous functions on [0,1]), $\langle f, g \rangle := \int_0^1 fg$
   inner product on $\mathbf{R}^{2\times 2}$: $\langle A, B \rangle = \mathrm{trace}(A^T B)$
   inner product on $\mathbf{R}^{2\times 2}$: $\langle A, B \rangle = a_{11}b_{11} + 2a_{12}b_{12} + 3a_{21}b_{21} + 4a_{22}b_{22}$
3. **fact**: every inner product on $\mathbf{R}^n$ is $\langle u, v \rangle = u^T A v$ where A is a symmetric (therefore diagonalizable) matrix with positive eigenvalues and $a_{ij} = \langle e_i, e_j \rangle$
4. **length (norm)**: $\|v\| = \sqrt{\langle v, v \rangle}$

5. **properties**:
$\|v\| \geq 0$
$\|v\| = 0$ iff $v = \underline{0}$
$\|\alpha v\| = |\alpha| \cdot \|\alpha v\|$
$\|u + v\| \leq \|u\| + \|v\|$
6. **unit vector**: $\|v\| = 1$
7. **unit vector in the direction of** v: $\frac{v}{\|v\|}$
8. **distance**: $d(u, v) = \|u - v\|$
9. **angle**: $\angle(u, v) = \arccos \frac{\langle u,v \rangle}{\|u\| \cdot \|v\|}$
10. **orthogonal**: $u \perp v$ iff $\angle(u, v) = \pi/2$ iff $\langle u, v \rangle = 0$
11. $S = \{v_1, \ldots, v_n\}$ **orthogonal**: $v_i \perp v_j$ for all i, j
12. **fact**: nonzero orthogonal vectors are independent
13. $S = \{v_1, \ldots, v_n\}$ **orthonormal**: S is orthogonal and $\|v_i\| = 1$ for all i
14. **Triangle inequality**:
$d(u, v) \leq d(u, w) + d(w, v)$
15. **orthogonal complement**: $W^\perp = \{v \in V \mid v \perp w \text{ for all } w \in W\}$, W is subspace of V
16. **properties**: W is subspace of $\mathbf{R}^n$
$W^\perp$ is a subspace
$W \cap W^\perp = \{\underline{0}\}$
$W = \text{span}(S)$, $u \perp s_i$ for all i implies $u \in W^\perp$
$(\text{Row}A)^\perp = \text{Null}A$
$\dim W + \dim W^\perp = n$
(basis of W) $\cup$ (basis of $W^\perp$) is basis of $\mathbf{R}^n$
$(W^\perp)^\perp = W$
17. **Pythagorean theorem**: $u \perp v$ implies $\|u + v\| = \|u\| + \|v\|$
18. **Cauchy-Schwartz inequality**: $|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$

17. Orthogonal bases and Gram-Schmidt algorithm

1. **notation**: $\{v_1, \ldots, v_n\}$ orthogonal basis, $\{b_1, \ldots, b_n\}$ orthonormal basis for a subspace W of V, $p \in V$
2. **orthogonal projection**: $\text{proj}_W v = \sum_{i=1}^n \frac{\langle p, v_i \rangle}{\langle v_i, v_i \rangle} v_i \in W$
3. **Gram-Schmidt algorithm**: for finding an orthogonal basis $\{b_1, \ldots, b_n\}$ for $\text{span}\{v_1, \ldots, v_n\}$
(i) make $\{v_1, \ldots, v_n\}$ independent if necessary
(ii) let $u_1 = v_1$
(iii) inductively let $u_{i+1} = v_{i+1} - \text{proj}_{\text{span}\{u_1, \ldots, u_i\}} v_{i+1} = v_{i+1} - \sum_{j=1}^i \frac{\langle v_{i+1}, u_i \rangle}{\langle u_i, u_i \rangle} u_i$
4. **fact**: $W = \text{Col}(A)$, $A\beta = \text{proj}_W y$ iff $A^T A \beta = A^T y$

18. Least square solution and linear regression

1. **fact**: if W subspace of V, $w \in W$, $y \in V$ then $\|y - w\|$ is minimum when $w = \text{proj}_W(y)$
2. **fact**: $W = \text{Col}(A)$, $\|y - A\beta\|$ is minimum iff $A^T A \beta = A^T y$
3. **least square regression line** $ax + b$: data $\{(x_i, y_i) \mid i = 1, \ldots, n\}$
$A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$, $\beta = \begin{pmatrix} b \\ a \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, β makes $\|A\beta - y\|$ minimum, that is, $A^T A \beta = A^T y$
$ax + b = \text{proj}_{\text{Col(A)}}(y)$