6.441 Recitation – Background Materials Review

April 4, 2003

Outline

- Some linear algebra review
  - Special matrices
  - Eigenvalues
  - Unitary decomposition
  - Singular values decomposition

- Gaussian random variables and random vectors
  - Motivation: Central Limit theorem
  - Gaussian random variables
  - Gaussian random vectors

Vectors and Matrices

- $\mathcal{R}^n$: $\mathcal{R}$-vector space, $\mathcal{C}^n$: $\mathcal{C}$-vector space

  with inner product $\langle x, y \rangle = \sum_i \bar{x}_i y_i = \bar{x}^T y$,
  $x_{n\times 1}, y_{n\times 1} \in \mathcal{C}^n$.
  Properties:
  1) $\overline{\langle x, y \rangle} = \langle y, x \rangle$
  2) $\langle \alpha x, y \rangle = \bar{\alpha}\langle x, y \rangle$
  3) $\langle x, \alpha y \rangle = \alpha\langle x, y \rangle$
  4) $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$
  5) $\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle$
  6) $\|x\|_2^2 \equiv \langle x, x \rangle = \sum_i |x_i|^2 > 0, \forall x \neq 0$.

  Basis $\{v_1, \ldots, v_n\}$ orthonormal if $\langle v_i, v_j \rangle =$ 0 when $i \neq j$ and 1 when $i = j$.

- Want to study: $\mathcal{C}$-linear transformations $L : \mathcal{C}^n \to \mathcal{C}^n$,
  Matrix representation: $L(x)_{n\times 1} = A_{n\times n} x_{n\times 1}$
  matrix : over $\mathcal{C}$ by default
  Can add and multiply (i.e compose)

Some special matrices

- Identity matrix I

- Real matrices: entries in $\mathcal{R}$

- Diagonal matrices

- Invertible matrices: if $\exists B$ s.t. $AB = I$, $A^{-1} \equiv B$.
  Equivalent definitions: 1) $Ax \neq 0$ for each $x \neq 0$ 2) $det A \neq 0$
  Invertible ones closed under multiplication.
  Non-invertible: singular

- Normal matrices: $\bar{A}^T A = A\bar{A}^T$.

- Hermitian matrices: if $\bar{A}^T = A$.
  Equivalent definitions: $\langle Ax, y\rangle = \langle x, Ay\rangle$.

Real symmetric matrices: if $A^T = A$, i.e. if Real and Hermitian

- Unitary Matrices: $\bar{A}^T A = I$.
  Equivalent definitions:

  1) preserves inner product , i.e. $\langle Ax, Ay \rangle = \langle x, y \rangle$
  2) maps orthonormal basis to orthonormal basis.
  3) columns form an orthonormal basis

  Orthogonal matrices: Real and Unitary

  Example: 2×2 rotation

$$\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

- Positive (definite): Hermitian and $\bar{x}^T A x \geq 0$ (> 0) for each x ($x \neq 0$) . Example:

  $$\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

  $a, b \in \mathcal{R}$ ($\neq 0$).

  We will see soon that: A positive $\Leftrightarrow \exists\ B$ s.t. $A = B\bar{B}^T$.

Eigenvalues and Eigenvectors

- Let A be an $n \times n$ matrix, $\lambda \in \mathcal{C}$, $v \neq 0 \in \mathcal{C}^n$. Then (λ, v) is an eigenvalue-eigenvector pair of A if $Av = \lambda v$.
  $\Leftrightarrow A - \lambda I$ singular $\Leftrightarrow det(A - \lambda I) = 0$

- Notes:
  1) $p(s) = det(A - sI)$ is a degree n-polynomial in s over $\mathcal{C}$, so $p(s) = \prod_{i=1}^{n}(s - \lambda_i)$
  2) λ_i s not necessarily distinct
  3) If A real, λ_i s can be complex

- If M is invertible and A arbitrary, Eigen values of A and MAM^{-1} are the same with the same multiplicities: If $\lambda v = Av$, then $(MAM^{-1})(Mv) = MAv = \lambda(Mv)$.
  Change of basis.

Special matrices in terms of eigenvalues

- Invertible $\Leftrightarrow \lambda_i \neq 0, \forall i$.
  Note : $detA = \prod_i \lambda_i$
- Hermitian $\Leftrightarrow$ normal and $\lambda_i \in \mathcal{R}, \forall i$.
  pf of $\Rightarrow$: $\bar{\lambda}\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle$.
- Unitary $\Leftrightarrow$ normal and $|\lambda_i| = 1, \forall i$.
  pf of $\Rightarrow$: $\lambda\bar{\lambda}\langle v, v\rangle = \langle \lambda v, \lambda v\rangle = \langle Av, Av\rangle = \langle v, v\rangle$.
  Rotation matrix example: $\lambda_1 = e^{i\phi}$, $\lambda_2 = e^{-i\phi}$
- Positive (definite) $\Leftrightarrow$ Hermitian and $\lambda_i \geq 0$ (> 0), $\forall i$.
  pf of $\Rightarrow$: $\lambda\langle v, v\rangle = \langle v, Av\rangle \geq 0$.
- Other directions follow form Unitary decomposition coming next.

Unitary Decomposition

- If A is normal then $\exists U$ unitary and D diagonal s.t. $A = UD\bar{U}^T$.
- Moreover if A is also real then so is U, i.e. U is orthogonal.
- eigenvalues on the diagonal.
- Corollary: A positive $\Leftrightarrow \exists\ B$ s.t. $A = B\bar{B}^T$
- Exercise: read the proof in your favorite book.

Singular Values Decomposition.

- Let A be any $n \times n$ matrix. $\bar{A}^T A$ is $n \times n$ hermitian.
  Also positive : $x\bar{A}^T A x^T = x\bar{A}^T \overline{(x\bar{A}^T)}^T \geq 0$
  So eigenvalues ≥ 0.
- Singular values of $A \equiv$ non-negative square-roots of eigenvalues of $A\bar{A}^T$.
- Singular Values Decomposition (SVD): $\exists$ U Unitary, V Unitary, s.t. $A = US\bar{V}^T$, where S is diagonal and positive.
  More over if A is also real, then so are U and V, i.e. U and V are orthogonal.
- Check: $\bar{A}^T A = VS\bar{U}^T US\bar{V}^T = VS^2\bar{V}^T$.
- Note: If A is positive , then Unitary decomposition is the same as the SVD.

- Exercise: prove SVD using Unitary decomposition of $\bar{A}^T A$.
- Exercise: Order the singular values of A : $\sigma_1 \geq \ldots \geq \sigma_n$. Show that

  1) $\max_{\|x\|_2=1} \|Ax\|_2 = \sigma_1$

  2) $\min_{\|x\|_2=1} \|Ax\|_2 = \sigma_n$

  3) $\sum_{i,j} A^2_{i,i} = \sum_i \sigma_i^2$

  Note: SVD works also for non-square matrices

Gaussian random variables and vectors

Motivation: Central Limit Theorem

- A simple form of the CLT: Let $Z_1, Z_2, \ldots$ be independent random variables s.t. $\exists c_1, c_2 > 0$ s.t. $c_1 < var(Z_i) < c_2, \forall i$.Let $a < b$. Then

$$\lim_{n \to \infty} Pr[a < \frac{\sum_{i=1}^n Z_i - E\sum_{i=1}^n Z_i}{var(\sum_{i=1}^n Z_i)} < b]$$

$$= \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx.$$

- Standard Gaussian random variable PDF: $p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.
- Properties: $EX = 0, Var(X) = 1$
  $\mathrm{E}e^{-iyX} = e^{-\frac{y^2}{2}}, \forall y \in \mathcal{C}$
  i.e. $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{-iyx} dx = \sqrt{2\pi} e^{-\frac{y^2}{2}}$.

Guassian random variables

- Y a guassian random variable if $\exists \sigma, m \in \mathcal{R}$ s.t. $Y = \sigma X + m$, X unit guassian.
- Thus mean $EY = m$, Variance $Var(Y) = \sigma^2$, and PDF

$$p_Y(y) = \frac{1}{|\sigma|} p_X\left(\frac{y-m}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

  Also $Ee^{-iyX} = e^{-iym} e^{-\frac{y^2\delta^2}{2}}$.
- P_Y is uniquely specified by σ and μ. Notation: $Y \sim N(m, \sigma^2)$.

Linear combinations of Gaussian random variables

- Let $Y \sim N(m, \sigma^2)$, then
  - $Y + b \sim N(m + b, \sigma^2)$ since $p_{Y+b}(z) = p_Y(z - b)$
  - $aY \sim N(m, (a\sigma)^2)$ since $p_{aY}(z) = \frac{1}{|a|} p_Y(\frac{z}{a})$
- Let $Y_1 \sim N(0, \sigma_1^2)$, $Y_2 \sim N(0, \sigma_2^2)$, then $Y_1 + Y_2 \sim N(0, \sigma_1^2 + \sigma_2^2)$, since

  $$Ee^{-iy(Y_1+Y_2)} = Ee^{-iyY_1} Ee^{-iyY_2} = e^{-\frac{1}{2}y^2(\sigma_1^2+\sigma_2^2)}$$

  for each s, and since $f(x) \mapsto \widehat{f}(y) = \int_{-\infty}^{\infty} f(x) e^{-iyx} dx$ ...
- Thus $Y_i \sim N(m_i, \sigma_i^2)$, $i = 1, \ldots, n$, independent $\Rightarrow \sum_i a_i Y_i + b \sim N(\sum_i a_i m_i + b, \sum_i a_i^2 \sigma_i^2)$.

Nonedegenerate Gaussian random vectors

- Y = $(\mathsf{Y}_1, \ldots, Y_n)^T$ Nondegenerate Guassian random vector if $\exists\ Q$ invertible real $n \times n$ matrix, $X = (X_1, \ldots, X_n)^T$, $X_i \sim N(0,1)$ indepent, and $m \in \mathcal{R}^n$ s.t. $Y = QX + m$.
- Thus any $\mathcal{R}$-linear combinartion of $Y_1, \ldots, Y_n$ is again a Guassian random variable

Properties of nondegenerate Guassian random vectors

- Mean: $EY = m$
- Covariance matrix:

$$\begin{aligned} Var(Y) &\equiv E(Y-m)(E-m)^T = EQX(QX)^T \\ &= QEXX^TQ^T = QQ^T \equiv \Lambda \end{aligned}$$

- Covariance matrix Λ strictly positive real matrix: In general for any random vector whose covariance matrix exists

$$\begin{aligned} a^T Var(Y)a &= Ea^T(Y-m)(Y-m)^T a \\ &= E(a^T(Y-m))^2 \geq 0, \end{aligned}$$

and here $\Lambda = QQ^T$ is also nonsingular.

- PDF. Change of variable formula *

$$p_Y(y) \;=\; J(y)P_X(Q^{-1}(Y-m))$$

where

$$J(y) = \left| det \left(\frac{\partial (Q^{-1}y)_j}{\partial y_i} \right)_{i,j} \right| = |detQ^{-1}| = \frac{1}{\sqrt{det\Lambda}},$$

and

$$P_X(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}x^T x}.$$

Thus

$$p_Y(y) = \frac{1}{\sqrt{det\Lambda}(2\pi)^{n/2}} e^{-\frac{1}{2}(y-m)^T \Lambda^{-1}(y-m)}.$$

- Uniquely determined by m and Λ

*

$$\int_{g(T)} f(x)dx_1 \ldots dx_n = \int_T f(g(y))|\text{det Jac } g(y)|dy_1 \ldots dy_n$$

when the first integral exists, and $g : \mathcal{R}^n \to \mathcal{R}^n$ is one-to-one, has continuous first-order partial derivatives, and Jac($g(y)$) is nonsingular for all y in T.

General Gaussian random vectors

- General possibly degenerate Gaussian random vector: allow Q to be non-invertible, or in general an $n \times m$ matrix

- Equivalent definition: If any $\mathcal{R}$-linear combination of $Y_1, \ldots, Y_n$ is a Gaussian random variable.
- Again uniquely determined by mean and covariance matrix
- Marginally Gaussian not sufficient. Example:

$$P_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} + w(y) \right),$$

where $w : \mathcal{R} \to \mathcal{R}$ s.t. $w(y) = -w(-y)$, $|w(y)| \leq \frac{1}{\sqrt{2\pi e}}$, $\forall y$, and w zeros outside $[-1, 1]$.