Hamming Codes

- messages: vectors in $\mathbb{F}_2^n$

$$\underset{\text{message}}{M} \xrightarrow{\text{encode}} \underset{\substack{\text{(valid)}\\ \text{codeword}}}{C_A} \xrightarrow{\text{send}} C_B \xrightarrow{\text{check \& correct}} C_A \xrightarrow{\text{decode}} M$$

Alice (G) — Bob (H)

↳ first (M length) elements of C_A.

- assumption: 0 or 1 errors per message
- redundancy = $\dfrac{\#\text{ bits in } C_A}{\#\text{ bits in } M}$
- for all $m \geq 2$, there is a $(\underbrace{2^m-1}_{\text{codeword length}}, \underbrace{2^m-m-1}_{\text{message length}})$ Hamming code
- $\underset{3\times 7}{H}$: parity check matrix, $HC_B = H_{*i} \Rightarrow C_B$ has error in i^{th} component. $H = (F \mid I_3)$
  - ↳ valid codewords for (7,4) Hamming code are the elements of $N(H)$
    - ↳ basis for $N(H)$: columns of $N = \left(\frac{I_4}{-F}\right) = G \longleftarrow$ "generator matrix" $GM = C_A \in N(H)$
  - → $C_A = C_B + e_i$, M = first 4 elements of C_A (rest are "parity check bits")

Determinants

- determinant = ratio of signed volume

"right hand rule" $\vec{Ae_1}$, $\vec{Ae_3}$, $\vec{Ae_2}$ ↳ R(+)

  - ↳ sign of determinant: orientation (chirality) → + preserved, − : reversed
  - → $|\det A| = \dfrac{n\text{-dim vol. of } A(R)}{n\text{-dim vol. of } R.}$

- Axiom 1: $\det I = 1$

Axiom 2: row swaps switch sign of det.

Axiom 3: det is linear as a function of each row: $\det\begin{pmatrix} \cdots \\ a^T + b^T \\ \cdots \end{pmatrix} = \det\begin{pmatrix} \cdots \\ a^T \\ \cdots \end{pmatrix} + \det\begin{pmatrix} \cdots \\ b^T \\ \cdots \end{pmatrix}$

Proposition 4: if A has 2 equal rows, $\det A = 0$ (Ax. 2)

Proposition 5: adding a multiple of one row to another row leaves det unchanged (Ax. 3 + Prop. 4)

Prop. 6: if A has a row of zeros, $\det A = 0$ (Prop. 5 + Prop. 4)

Prop. 7: if A is UT/LT, then $\det A$ = product of entries on diagonal (Prop. 5)

Prop. 8: A is $n\times n$ and not invertible $\Leftrightarrow \det A = 0$ ($|\det A| = |\det U|$)

↳ Prop.: if A is invertible, $\det A^{-1} = 1/\det A$

Prop. 9: if A, B are $n\times n$, $\det AB = \det A \det B$

↳ Prop: if P is a permutation matrix, $\det P = \pm 1$ ← $\det P = (-1)^x$, $x = \#$ of row swaps from P to I

Prop. 10: if A is $n\times n$, $\det A^T = \det A$

- Permutation formula: $\sum_{\substack{n \times n \\ \text{permutation matrices}}}$ (det P)(product of entries in A where P has 1's)
  - ← $n!$ terms. ↳ only consider those w/ nonzero entry in each col. ← else, n^n determinants.

- Cofactor formula: $C_{ij} = (-1)^{i+j} * \det (\;) \leftarrow (n-1)\times(n-1)$ matrix.

  ↳ $\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \dots + a_{in}C_{in}$

  → $A^{-1} = \frac{1}{\det A} C^T$, $C = \begin{pmatrix} c_{11} & c_{12} & \cdots \\ \vdots & \ddots & \\ & & c_{nn} \end{pmatrix}$

- for 2x2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\det A = ad - bc$, $A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$

- Cramer's rule: if $A\vec{x} = \vec{b}$, then $x_i = \frac{\det B_i}{\det A}$, where $B_i = A$ with col i replaced by $\vec{b}$

Eigenvalues & Eigenvectors

- $A\vec{x} = \lambda\vec{x} \Rightarrow (A - \lambda I)\vec{x} = 0$, want to find solution for $\vec{x} \neq \vec{0}$ $\Rightarrow \vec{x} \in N(\underline{A - \lambda I})$ $\Leftrightarrow$ $\det(A - \lambda I) = 0$
  - not invertible.
  - "characteristic polynomial" for A.

- thm: a polynomial with degree n has at most n real roots.
  ↳ exactly n complex roots, counted w/ multiplicity.

- Arithmetic multiplicity: multiplicity of λ as a root of $\det(A - \lambda I)$ $\quad$ * $1 \le GM \le AM$.

  Geometric multiplicity: # of independent eigenvectors w/ eigenvalue λ = $\dim N(A - \lambda I)$

- A, C are similar if $A = BCB^{-1}$

  ↳ if A, C are similar, they have the same eigenvalues. & if $\vec{x}$ is an e-vec for C, $B\vec{x}$ is e-vec for A.

- A is diagonalizable if it is similar to a diagonal matrix

  ↳ $n \times n$ matrix A is diagonalizable $\Leftrightarrow$ A has n independent e-vecs (GM = AM for all λ)

  ↳ $A = X \Lambda X^{-1}$, $X = (\vec{x}_1 | \vec{x}_2 | \dots | \vec{x}_n)$, $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \;\lambda_n \end{pmatrix}$

  → A and Λ are connected by change of basis matrix X: $X\vec{e}_i = \vec{x}_i$, $X^{-1}\vec{x}_i = \vec{e}_i$
  - std. basis → eigenbasis.

  ↳ A, Λ are the same transformation wrt different bases. → $\det A = \det \Lambda$

- $n \times n$ matrix M is a Markov matrix if $M_{ij} \ge 0$ for $1 \le i,j \le n$ $\wedge$ $\sum_{i=1}^{n} = 1$ for $1 \le j \le n$ (col. sum = 1)

  ↳ 'positive' Markov matrix if $M_{ij} > 0$. $\quad$ ↳ $M^T \mathbb{1} = \mathbb{1}$

- A Markov chain has states and transition probabilities between states
  - ≥ 0, sum to 1, only depend on "prior" state (memoryless).

  ↳ M_{ij} = transition prob. from state j (prior) to state i (subsequent) $\longrightarrow$ $\vec{u}_{i+1} = M\vec{u}_i$

- if M is a diagonalizable Markov matrix with e-vec $\vec{x}_1$ and $\lambda_1 = 1$, and all other $|\lambda| < 1$, then

→ $\vec{x}_1$ is an attracting steady state: $M\vec{x}_1 = \vec{x}_1$, $M^k\vec{v} \xrightarrow{n\to\infty} c_1\vec{x}_1$, $\vec{v} = c_1\vec{x}_1 + c_2\vec{x}_2 + \ldots$

→ positive Markov matrices: $\lambda_1 = 1$, $|\lambda_i| < 1 \Rightarrow$ attracting steady state $\vec{x}_1$

↳ $\vec{v}$, $c\vec{x}_1$ have same component sum.

- A, A^T have the same eigenvalues

↳ every Markov matrix has eigenvalue $\lambda = 1$ $(M^T \mathbb{1} = 1 \cdot \mathbb{1})$

→ sum of the components of $M\vec{x}$ = sum of components of $\vec{x}$

→ if A, B are $n \times n$ Markov matrices, then AB is Markov

→ if $k \geq 1$, M Markov $\Rightarrow M^k$ Markov.

→ if λ is an eigenvalue for an $n \times n$ Markov matrix M, then $|\lambda| \leq 1$

- Spectral Theorem: if S is a real, symmetric $n \times n$ matrix, then $S = Q \Lambda Q^{-1} = Q \Lambda Q^T$

↳ $\underline{Q}$ is orthonormal basis of e-vectors, Λ is diagonal matrix of real e-values, S is "orthogonally diagonalizable".

↳ rotation if $\det = 1$, reflection if $\det = -1$.

- $p(\bar{x}) = \overline{p(x)}$, roots of p come in conjugate pairs

↳ real, symmetric matrix S has only real eigenvalues

→ real, symmetric matrix S has orthogonal eigenvectors

} proof of spectral thm.

- $n \times n$ symmetric matrix S is positive definite if — positive semi-definite

(i) all n pivots > 0 — ≥ 0

(ii) all upper left determinants are > 0 — ≥ 0

(iv) ⇒ (iii) ⇒ (v) ⇒ (iv)

(iii) all n eigenvalues are > 0 — ≥ 0

(iv) for all $\vec{x} \neq \vec{0}$, $\vec{x}^T S \vec{x} > 0$ — ≥ 0

(v) there is an $m \times n$ matrix A with indp. cols. s.t. $S = A^T A$ — … A such that $S = A^T A$ (symmetric)

Singular Value Decomposition

- $\overset{m\times n}{A} = \overset{m\times m}{U}\overset{m\times n}{\Sigma}\overset{n\times n}{V^T} = \overset{m\times r}{U_r}\overset{r\times r}{\Sigma_r}\overset{r\times n}{V_r^T}$ ← V_r is $n\times r$

↳ U, V are orthogonal matrices, Σ_r diagonal ← singular values, >0

→ for $1 \le i \le r$, $A\vec{v}_i = \sigma_i \vec{u}_i$ → $\{\vec{v}_1, \dots, \vec{v}_r\}$ form orthonormal basis for $C(A^T)$,

$\{\vec{u}_1, \dots, \vec{u}_r\}$ form orthonormal basis for $C(A)$

for $i > r$, $A\vec{v}_i = \vec{0} = \sigma_i \vec{u}_i$ → $\{\vec{v}_{r+1}, \dots, \vec{v}_n\}$ form orthonormal basis for $N(A)$

$\{\vec{u}_{r+1}, \dots, \vec{u}_m\}$ form orthonormal basis for $N(A^T)$

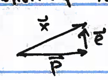
- $\vec{v}_i$'s are the e-vecs of A^TA, $\vec{u}$'s are the e-vecs of AA^T

↳ σ_i's are the square root of the positive e-values of A^TA & AA^T

↳ A^TA and AA^T have no neg e-values, the same pos. e-values, and the e-value 0 multiplicity $n-r$ / $m-r$

- $\vec{v}$'s $\xrightarrow{V^T}$ $\vec{e}$'s $\xrightarrow{\Sigma}$ stretch $\xrightarrow{U}$ $\vec{u}$'s (reflection)

① Best-fit Subspace: given $\vec{x}_1, \dots, \vec{x}_m \in \mathbb{R}^n$, find orthonormal set of vectors $\underbrace{\{\vec{v}_1, \dots, \vec{v}_k\}}_{\text{first } k \text{ singular vectors}}$ spanning $\underbrace{W \subseteq \mathbb{R}^n}_{k\text{-dim}}$ s.t.

↳ $\sum_{i=1}^{m} \|\vec{e}_i\|^2$ is minimized ⟺ $\sum_{i=1}^{m} \|\vec{p}_i\|^2$ is maximized 

↳ i.e. WTF $\{\vec{v}_1, \dots, \vec{v}_k\}$ s.t. $\vec{x}_i \simeq \vec{p}_i = \mathrm{proj}_{\vec{v}_1}\vec{x}_i + \dots + \mathrm{proj}_{\vec{v}_k}\vec{x}_i = (\vec{x}_i^T\vec{v}_1)\vec{v}_1 + \dots + (\vec{x}_i^T\vec{v}_k)\vec{v}_k$

→ if Λ is a pos. semi-definite diagonal matrix where $\lambda_1 \ge \lambda_2 \ge \dots \ge 0$

↳ the unit vector $\vec{y}$ maximizing $\vec{y}^T\Lambda\vec{y}$ is $\vec{y} = \vec{e}_1$

for any

→ $A = U\Sigma V^T$ ⇒ $A^TA = Q\Lambda Q^T$ where $Q = V$, $\Lambda = \Sigma^T\Sigma$ ⇒ unit vector $\vec{v}$ maximizes $\|A\vec{v}\|$ when $\vec{v} = \vec{v}_1$

↳ $\vec{v}_1$ is the first singular vector of $A = \begin{pmatrix} \vec{x}_1^T \\ \vdots \\ \vec{x}_m^T \end{pmatrix}$

↳ line of perpendicular best fit (1-D subspace of best fit) → unlabelled data

→ least squares line of best fit - labelled data.

② Low rank matrix approximation: $AB = A_{*1}B_{1*}^T + \dots + A_{*n}B_{n*}^T$ (rows of AB are lin. comb of rows of B)

↳ block multiplication: $\underset{\substack{1\times n \text{ block matrix} \\ m\times 1 \text{ blocks}}}{\overset{m\times n}{(\,|\,|\,|\,|\,|\,)}}\underset{\substack{n\times 1 \text{ block matrix} \\ 1\times p \text{ blocks}}}{\overset{n\times p}{(\equiv)}}$ = 1×1 block matrix of $m\times p$ blocks

each row is a lin comb of $\vec{v}_1, \dots, \vec{v}_r$ → $A = \vec{u}_1\vec{v}_1^T + \dots + \vec{u}_r\vec{v}_r^T$ ↗ rank r matrix = sum of r rank 1 matrices

→ $A_k = U_k\Sigma_kV_k^T$ → $\vec{x}_i^T = \underbrace{\sigma_1 u_{i1}\vec{v}_1^T + \dots + \sigma_k u_{ik}\vec{v}_k^T}_{\text{lin comb of } \{\vec{v}_1, \dots, \vec{v}_k\}} + \underbrace{\dots + \sigma_r u_{ir}\vec{v}_r^T}_{\text{lin comb of } \{\vec{v}_{k+1}, \dots, \vec{v}_r\}}$ → $\vec{e}_i \in U^\perp$ } i^{th} row of A_k: $\vec{p}_i$

$\vec{p}_i \in U = \mathrm{span}\{\vec{v}_1, \dots, \vec{v}_r\}$