

Exam 1

1. Systems A and B consist of four linear equations in four unknowns modulo 2.

$$\begin{array}{ccccccc} x_1 & + & x_2 & = & 0 \\ + & & + & & \\ y_1 & + & y_2 & = & 0 \\ = & & = & & \\ 1 & & 1 & & \end{array}$$

system A

$$\begin{array}{ccccccc} x_1 & + & x_2 & = & 1 \\ + & & + & & \\ y_1 & + & y_2 & = & 0 \\ = & & = & & \\ 1 & & 1 & & \end{array}$$

system B

- (a) Solve system A using modular Gauss-Jordan elimination. If there are multiple solutions, specify the free variables and the assignment to the remaining variables in terms of the free variables.

Solution: Keeping the equation $x_1 + x_2 = 0$ and eliminating x_1 from the others gives the system $x_2 + y_1 = 1$, $x_2 + y_2 = 1$, $y_1 + y_2 = 0$. Now we keep $x_2 + y_1 = 1$ and eliminate x_2 to obtain $y_1 + y_2 = 0$ twice. Eliminating y_1 leaves y_2 as the free variable. Substituting back gives the solution $y_1 = y_2$, $x_2 = y_2 + 1$, $x_1 = y_2 + 1$. (As y_2 can be either 0 or 1 there are two solutions for (x_1, x_2, y_1, y_2) : $(1, 1, 0, 0)$ and $(0, 0, 1, 1)$.)

- (b) In system B the value in the first equation was flipped. Produce a contradictory linear combination for system B . You may use any method you like.

Solution: The contradictory linear combination is the sum of all the equations. As each variable appears twice the left-hand side vanishes modulo two. The right hand side equals $1 + 0 + 1 + 1 = 1$.

2. You want to apply Gradient Descent to the system

$$\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The eigenvalues of the matrix are $\lambda_1 = 4$ and $\lambda_2 = -2$ with corresponding eigenvectors $\mathbf{v}_1 = (1/\sqrt{2}, 1/\sqrt{2})$, $\mathbf{v}_2 = (1/\sqrt{2}, -1/\sqrt{2})$.

- (a) What is the maximum rate ρ^* below which convergence to the unique solution $(1/4, 1/4)$ is guaranteed (for any initialisation)?

Solution: ρ^* is the inverse square of the spectral norm of A . The spectral norm of a symmetric matrix is the largest eigenvalue in absolute value, which is 16 in this case. Therefore $\rho^* = 1/16$.

- (b) You run Gradient Descent with rate $\rho = 0.01$ and initialisation $(x, y) = (1, 0)$. The distance between the state (x_t, y_t) at time t and the solution $(1/4, 1/4)$ is $\Theta(b^t)$ for some number b . Find b .

Solution: $b = 0.92$. Let $\mathbf{x}^* = (1/4, 1/4)$ and $\mathbf{x} = (x, y)$ be the state of the algorithm at some point. The difference $\mathbf{x} - \mathbf{x}^*$ is multiplied by the matrix $B = I - 2\rho AA^T$ at every step of the algorithm. As A is symmetric, AA^T has the same eigenvectors as A^2 and therefore A , and so does $B = I - 2\rho AA^T$. The associated eigenvalues are $1 - 2\rho\lambda_1^2 = 0.68$ and $1 - 2\rho\lambda_2^2 = 0.92$. The initial state has nonzero projection in the direction of $\mathbf{v}_2$, so the latter number governs the rate of convergence.

3. Calculate the Fourier transforms of the following functions. You may use any method you like. $+$ and $-$ are shorthand for the numbers $+1$ and -1 .

$$(a) \frac{x_1 x_2}{f(x_1 x_2)} \left| \begin{array}{c|cccc} & ++ & -+ & +- & -- \\ \hline & 1 & 1 & -1 & -1 \end{array} \right.$$

Solution: This is the function x_2 . Its only nonzero Fourier coefficient is $\hat{f}(\{2\}) = 1$.

$$\frac{S}{\hat{f}(S)} \left| \begin{array}{c|cccc} & \emptyset & 1 & 2 & 12 \\ \hline & 0 & 0 & 1 & 0 \end{array} \right.$$

$$(b) \frac{x_1 x_2}{g(x_1 x_2)} \left| \begin{array}{c|cccc} & ++ & -+ & +- & -- \\ \hline & 0 & 1 & 1 & 1 \end{array} \right.$$

(**Hint:** What is $1 - g$?)

Solution: $1 - g$ is the point function point $_{++}$. Its polynomial representation is $(1 + x_1)/2 \cdot (1 + x_2/2) = \frac{1}{4} + \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_1x_2$. Therefore $g = \frac{3}{4} - \frac{1}{4}x_1 - \frac{1}{4}x_2 - \frac{1}{4}x_1x_2$.

$$\frac{S}{\hat{g}(S)} \left| \begin{array}{c|cccc} & \emptyset & 1 & 2 & 12 \\ \hline & 3/4 & -1/4 & -1/4 & -1/4 \end{array} \right.$$

$$(c) h(x_1 x_2 x_3) = \begin{cases} f(x_1 x_2), & \text{if } x_3 = -1 \\ g(x_1 x_2), & \text{if } x_3 = +1. \end{cases}$$

(**Hint:** Apply one of the steps of the Fast Fourier-Walsh algorithm.)

Solution:

$$\begin{aligned} h(x_1 x_2 x_3) &= \frac{1 - x_3}{2} \cdot f + \frac{1 + x_3}{2} \cdot g \\ &= \frac{1}{2}(g + f) + \frac{1}{2}(g - f)x_3 \\ &= \frac{1}{2}\left(\frac{3}{4} - \frac{1}{4}x_1 + \frac{3}{4}x_2 - \frac{1}{4}x_1x_2\right) + \frac{1}{2}\left(\frac{3}{4} - \frac{1}{4}x_1 - \frac{5}{4}x_2 - \frac{1}{4}x_1x_2\right)x_3 \end{aligned}$$

which gives the coefficient list

$$\frac{S}{\hat{g}(S)} \left| \begin{array}{c|cccccccc} & \emptyset & 1 & 2 & 3 & 12 & 13 & 23 & 123 \\ \hline & \frac{3}{8} & -\frac{1}{8} & \frac{3}{8} & \frac{3}{8} & -\frac{1}{8} & -\frac{1}{8} & -\frac{5}{8} & -\frac{1}{8} \end{array} \right.$$

Exam 2

1. Let S be the system of equations

$$\begin{aligned}x + y + z &= 0 \\x + 2y + 4z &= 0\end{aligned}$$

- (a) Apply Gauss-Jordan elimination to find a nonzero solution. Explain your steps.

Solution: After eliminating x we obtain $y + 3z = 0$. After eliminating y , z becomes a free variable. The solution found by Gauss-Jordan Elimination is $FREE = \{z\}$, $y = -3z$, $x = -y - z = 2z$. Setting z to any nonzero value yields a nonzero solution. For example, $z = 1$ gives $y = -3$ and $x = 2$.

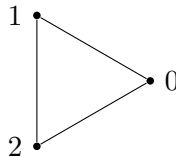
- (b) Now add the equation $x + 3y + 7z = 1$ to the system. Find a contradictory linear combination.

Solution: We seek coefficients a, b, c so that $a(x + y + z) + b(x + 2y + 4z) + c(x + 3y + 7z)$ is always zero, but $0a + 0b + 1c = c$ is not zero. We may set $c = 1$. For this it is sufficient to solve

$$\begin{aligned}a + b + 1 &= 0a + 2b + 3 & &= 0a + 4b + 7 = 0\end{aligned}$$

Subtracting the first two equations we get $b = -2$ and $a = 1$. Indeed, adding the first and third equation and subtracting the second gives zero on the left and one on the right.

2. Let A be the adjacency matrix of the 3-cycle (with zeros on the diagonal). The largest eigenvalue of A is $\lambda_1 = 2$. The corresponding eigenvector is $\mathbf{v}_1 = (1/\sqrt{3})(1, 1, 1)$.



- (a) Let $\mathbf{x}$ be any vector that is orthogonal to $\mathbf{v}_1$. Show that the state of Power Iteration on A initialized with $\mathbf{x}$ oscillates between $\mathbf{x}$ and $-\mathbf{x}$. Assume there are no precision errors.

Solution: Let $\mathbf{x} = (x_1, x_2, x_3)$. As it is orthogonal to $\mathbf{v}_1$, $x_1 + x_2 + x_3 = 0$. Applying A to $\mathbf{x}$ gives $A\mathbf{x} = (x_2 + x_3, x_3 + x_1, x_1 + x_2) = (-x_1, -x_2, -x_3) = -\mathbf{x}$ because subtracting $x_1 + x_2 + x_3$ from every coordinate does not change $\mathbf{x}$.

- (b) Use part (a) to deduce the other two eigenvalues λ_2 and λ_3 of A . What are they?

Solution: $\lambda_2 = \lambda_3 = -1$. As eigenvectors are mutually orthogonal, applying A to an eigenvector $\mathbf{v}$ other than $\mathbf{v}_1$ outputs $-\mathbf{v}$, so the corresponding eigenvalue must be -1 .

3. Apply the modular Fast Fourier transform to calculate the polynomial representation of

$$\begin{array}{c|cccc}x & 1 & i & -1 & -i \\ \hline f(x) & 0 & 1 & 2 & 3\end{array}$$

- (a) Calculate the list representations of f^+ and f^- , where $f(x) = f^+(x^2) + xf^-(x^2)$.

Solution: $f^+(x^2) = (f(x) + f(-x))/2$ has list representation $f(1) = 1, f(-1) = 2$. $f^-(x^2) = (f(x) - f(-x))/2x$ has list representation $f(1) = -1, f(-1) = i$.

- (b) Calculate the polynomial representations of f^+ and f^- over domain $\{-1, +1\}$.

Solution: g over domain $-1, 1$ has polynomial representation $g(y) = \frac{1}{2}(g(1) + g(-1)) + \frac{1}{2}(g(1) - g(-1)) \cdot y$. Applying this formula we get $f^+(y) = \frac{3}{2} - \frac{1}{2} \cdot y$ and $f^-(y) = (-\frac{1}{2} + \frac{1}{2}i) + (-\frac{1}{2} - \frac{1}{2}i) \cdot y$.

- (c) Calculate the polynomial representation of f .

Solution: $f(x) = f^+(x^2) + xf^-(x^2) = \frac{3}{2} + (-\frac{1}{2} + \frac{1}{2}i) \cdot x - \frac{1}{2} \cdot x^2 + (-\frac{1}{2} - \frac{1}{2}i) \cdot x^3$.

Exam 3

1. You apply gradient descent to the following (inconsistent) univariate linear system:

$$x = 0$$

$$x = 1$$

- (a) Write the sum-of-squares loss f . For which value of x^* is $f(x^*)$ minimized?

Solution: The sum-of-squares loss is $f(x) = x^2 + (x-1)^2$. Its gradient is $f'(x) = 2x + 2(x-1) = 2(2x-1)$. The loss is minimized when the gradient is zero, namely at $x^* = 1/2$.

- (b) Let x_t be the state after t steps of gradient descent with rate ρ . What is $(x_t - x^*)/(x_{t-1} - x^*)$?

Solution: It equals $1 - 4\rho$. The state evolution equation is $x_t = x_{t-1} - \rho f'(x_{t-1}) = x_{t-1} - 2\rho(2x_{t-1} - 1)$. Therefore $x_t - 1/2 = (x_{t-1} - 1/2) - 2\rho(2x_{t-1} - 1) = (1 - 4\rho)(x_{t-1} - 1/2)$.

- (c) Assuming $x_0 = 0$ and $\rho = 1/8$, what is x_{20} ?

Solution: By part (b),

$$x_{20} - 1/2 = (1 - 4\rho)^{20}(x_0 - 1/2) = (1/2)^{20}(-1/2) = -(1/2)^{21}.$$

2. You apply subspace iteration on the matrix

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

Your initial state is the standard basis $\mathbf{x}_1 = (1, 0, 0)$, $\mathbf{x}_2 = (0, 1, 0)$, $\mathbf{x}_3 = (0, 0, 1)$.

- (a) Ignoring $\mathbf{x}_3$ for now, show how $\mathbf{x}_1$ and $\mathbf{x}_2$ evolve in the first three rounds of subspace iteration.

Solution: Round 1: $A\mathbf{x}_1$ and $A\mathbf{x}_2$ are the first two columns of A . They are already orthogonal but not of unit norm. After orthogonalization the state is $\mathbf{x}_1 = (0, 1/\sqrt{5}, 2/\sqrt{5})$ and $\mathbf{x}_2 = (1, 0, 0)$.

Round 2: $A\mathbf{x}_1$ now becomes $(\sqrt{5}, 0, 0)$ and $A\mathbf{x}_2$ becomes $(0, 1, 2)$. After orthogonalization $\mathbf{x}_1 = (1, 0, 0)$ and $\mathbf{x}_2 = (0, 1/\sqrt{5}, 2/\sqrt{5})$. They switch places from round 2.

Round 3: They now switch places again, so the outcome is the same as after round 1. The state keeps iterating between $\mathbf{x}_1 = (0, 1/\sqrt{5}, 2/\sqrt{5})$, $\mathbf{x}_2 = (1, 0, 0)$ and $\mathbf{x}_1 = (1, 0, 0)$, $\mathbf{x}_2 = (0, 1/\sqrt{5}, 2/\sqrt{5})$.

- (b) Use part (b) to find two of the three eigenvalues of A . Explain your reasoning.

Solution: The two eigenvalues are $\sqrt{5}$ and $-\sqrt{5}$. As $A\mathbf{x}_1 = \sqrt{5}\mathbf{x}_2$ and $A\mathbf{x}_2 = \sqrt{5}\mathbf{x}_1$, $\mathbf{x}_1 + \mathbf{x}_2$ is an eigenvector with eigenvalue $\sqrt{5}$ and $\mathbf{x}_1 - \mathbf{x}_2$ is an eigenvector with eigenvalue $-\sqrt{5}$.

- (c) Prove that if, at any point during power iteration, $A\mathbf{x}_i$ and $A\mathbf{x}_j$ become linearly dependent (for some $i \neq j$), zero must be an eigenvalue of A .

Solution: if $A\mathbf{x}_i = cA\mathbf{x}_j$ then $A(\mathbf{x}_i - c\mathbf{x}_j)$ must be zero. But $\mathbf{x}_i$ and $\mathbf{x}_j$ are orthogonal so $\mathbf{x}_i - c\mathbf{x}_j$ is a nonzero eigenvector with eigenvalue zero.

- (d) Use part (c) to find the remaining eigenvalue of A . (**Hint:** What happens to $\mathbf{x}_3$?)

Solution: In round 1 already $A\mathbf{x}_2$ and $A\mathbf{x}_3$ are linearly dependent as they are both multiples of $(0, 1, 2)$. By part (c) zero is an eigenvalue of A .

3. The 2-bit addressing function $\text{addr}: \{+1, -1\}^6 \rightarrow \{+1, -1\}$ is given by

$$\text{addr}(x, y, z_{++}, z_{+-}, z_{-+}, z_{--}) = z_{xy}.$$

- (a) Express addr in terms of the four point functions $\text{point}_{ab}(x, y)$ and the four variables z_{ab} .

Solution: $\text{addr} = z_{++}\text{point}_{++}(x, y) + z_{+-}\text{point}_{+-}(x, y) + z_{-+}\text{point}_{-+}(x, y) + z_{--}\text{point}_{--}(x, y)$.

(b) Calculate $\widehat{\text{addr}}(\emptyset)$, $\widehat{\text{addr}}(\{1, 2, 3\})$ and $\widehat{\text{addr}}(\{1, 2, 3, 4\})$. (Use part (a).)

Solution: $\widehat{\text{addr}}(\emptyset) = 0$: All terms in the polynomial expansion of addr contain of the z_{ab} variables. The constant term must therefore be zero.

$\widehat{\text{addr}}(\{1, 2, 3\}) = 1/4$: This is the coefficient in front of the monomial xyz_{++} . This is the xy -coefficient in the Fourier expansion of $\text{point}_{++}(x, y)$. We calculated it as $1/4$ in class.

$\widehat{\text{addr}}(\{1, 2, 3, 4\}) = 0$: The degree of addr can be at most 3, as each of the point functions has degree 2. The term $xyz_{++}z_{+-}$ does not appear in the expansion.