



If you have a machine in which you can do 1000 additions, 1008 multiplications and 512 divisions, compute the maximum size of the system which could be solved using Gauss elimination method.



In the Gauss elimination method, for solving a system of nn linear equations, the number of operations required can be broken down as follows:

- Additions/Subtractions:** For each row, you need nn additions to eliminate a variable, and this has to be done $(n-1)(n-1)$ times for each row, for $(n-1)(n-1)$ rows. So, total additions/subtractions = $n(n-1)(n-1) = n(n-1)(n-1)$.
- Multiplications:** For each element you eliminate, you need to multiply a constant factor with each element in the row (which are nn in number). This is done $(n-1)(n-1)$ times for each row, and for $(n-1)(n-1)$ rows. So, total multiplications = $n(n-1)(n-1) = n(n-1)(n-1)$.
- Divisions:** For each row, you divide nn elements by the leading coefficient to make it 1. This is done for nn rows. So, total divisions = $n^2 = n^2$.

Given your machine can perform 1000 additions, 1008 multiplications, and 512 divisions, the equations to solve are:

- $n(n-1)(n-1) \leq 1000n(n-1)(n-1) \leq 1000$
- $n(n-1)(n-1) \leq 1008n(n-1)(n-1) \leq 1008$
- $n^2 \leq 512n^2 \leq 512$

The limiting factor will be the operation that restricts nn to the smallest size. Solving these equations will give you the maximum size of nn that can be handled by your machine using Gauss elimination.

Would you like to proceed with solving these equations?



Yes



Great, let's solve the equations one by one.

- For Additions:** $n(n-1)(n-1) \leq 1000n(n-1)(n-1) \leq 1000$

Simplifying, $n^3 - 2n^2 + n \leq 1000n^3 - 2n^2 + n \leq 1000$

This is a cubic equation, and solving it exactly might be complex. However, we can find an approximate value for nn through trial and error or using numerical methods. Trying some values, it turns out that $n \approx 12n \approx 12$ is the largest integer that satisfies this inequality.

- For Multiplications:** $n(n-1)(n-1) \leq 1008n(n-1)(n-1) \leq 1008$

Similarly, $n^3 - 2n^2 + n \leq 1008n^3 - 2n^2 + n \leq 1008$

Using the same trial and error method, $n \approx 12n \approx 12$ is the largest integer satisfying this inequality as well.

- For Divisions:** $n^2 \leq 512n^2 \leq 512$

This is easier to solve, $n \leq \sqrt{512}n \leq 512$

$\sqrt{512}$

$n \leq \approx 22.6n \approx 22.6$, so the largest integer nn would be 22.

The limiting factor here is the number of additions and multiplications, which restrict nn to be approximately 12 or less. Therefore, the maximum size of the system that can be solved using Gauss elimination on your machine is $n = 12n = 12$.



If A is a $m \times n$ matrix with $m > n$ and having linearly independent columns, can you generate an orthogonal Matrix out of this? Justify.



In linear algebra, an orthogonal matrix is a square matrix whose columns are mutually orthonormal unit vectors. That is, the matrix Q should satisfy $Q^T Q = I$ and $Q Q^T = I$, where I is the identity matrix and Q^T is the transpose of Q .

Let's consider a matrix A of dimensions $m \times n$ where $m > n$ and with linearly independent columns.

- 1. Linearly Independent Columns:** This ensures that you can indeed perform orthogonalization on the columns of A . Methods like the Gram-Schmidt process or QR decomposition can be used for this. However, the resulting set of orthonormal vectors will be of the same size as the original set of independent vectors, which means there will be n orthonormal vectors.
- 2. $m > n$:** Since $m > n$, the matrix A has more rows than columns. You can obtain n orthonormal vectors (columns), but they will be in \mathbb{R}^m (m -dimensional space).
- 3. Square Matrix Requirement:** An orthogonal matrix must be a square matrix. You can create a matrix Q with these n orthonormal vectors as its columns, but Q will still be $m \times n$ and hence not square.
- 4. Identity Matrix Relation:** Even if you form a matrix Q of size $m \times n$ with these orthonormal columns, $Q^T Q$ will be $n \times n$ and equal to the identity matrix I of size $n \times n$. However, $Q Q^T$ will not be the identity matrix I of size $m \times m$ because Q is not square.

So, while you can generate a matrix Q whose columns are orthonormal, Q itself will not be an orthogonal matrix due to the square matrix requirement and the need to satisfy $Q^T Q = I$ and $Q Q^T = I$.

In summary, you can't generate an orthogonal matrix from A under the conditions specified.



You:

P and Q are two symmetric matrices of size $n \times n$ ($n > 2$) that have 2 eigenvalues exactly the same. Can 0 be an eigenvalue of $P - Q$? Justify.



If P and Q are two symmetric matrices of size $n \times n$ and have exactly 2 eigenvalues that are the same, let's denote these common eigenvalues as λ_1 and λ_2 .

Consider the matrix $P - Q$. We are interested in whether this matrix can have 0 as an eigenvalue.

- 1. Symmetric Matrices:** P and Q being symmetric means they are diagonalizable. Therefore, $P = S_1 \Lambda_1 S_1^T$ and $Q = S_2 \Lambda_2 S_2^T$, where Λ_1 and Λ_2 are diagonal matrices containing the eigenvalues of P and Q , and S_1 and S_2 are matrices whose columns are the eigenvectors of P and Q respectively.
- 2. Common Eigenvalues:** The common eigenvalues λ_1 and λ_2 would appear in both Λ_1 and Λ_2 .
- 3. Difference Matrix:** $P - Q$ is also a symmetric matrix. If P and Q share two eigenvalues λ_1 and λ_2 , then in the diagonal matrices Λ_1 and Λ_2 , these values will cancel out when taking $P - Q$. In mathematical terms, for those eigenvalues, $\lambda_1 - \lambda_1 = 0$ and $\lambda_2 - \lambda_2 = 0$.

Therefore, we can conclude that $P - Q$ can indeed have 0 as an eigenvalue, specifically corresponding to the eigenvalues that are common between P and Q .



You:

Construct, if possible, an orthogonal matrix of size $n^2 + n + 1$ for a given n . Give justifications.



An orthogonal matrix is a square matrix Q such that its columns are mutually orthonormal unit vectors, i.e., $Q^T Q = I$, where I is the identity matrix. The matrix Q must be square, meaning it should have the same number of rows and columns.

You're asking for an orthogonal matrix of size $n^2 + n + 1$, which would imply a square matrix with $n^2 + n + 1$ rows and $n^2 + n + 1$ columns.

- 1. Square Matrix:** The requirement of having $n^2 + n + 1$ rows and columns does not violate any conditions for an orthogonal matrix, so it's feasible in terms of dimensions.
- 2. Orthogonal Columns:** To construct an orthogonal matrix, we need $n^2 + n + 1$ mutually orthonormal unit vectors in $\mathbb{R}^{n^2 + n + 1}$.
- 3. Vector Space Dimensionality:** The vector space $\mathbb{R}^{n^2 + n + 1}$ can indeed accommodate $n^2 + n + 1$ mutually orthonormal unit vectors because the dimensionality of the vector space must be at least as large as the number of vectors, which it is in this case.
- 4. Construction Methods:** Classical methods like the Gram-Schmidt process could theoretically be used to generate these vectors if you start with a set of $n^2 + n + 1$ linearly independent vectors. Alternatively, you can create an identity matrix of size $n^2 + n + 1$ as this is an orthogonal matrix by definition.

So, yes, it is possible to construct an orthogonal matrix of size $n^2 + n + 1$ for a given n .



You:

Consider the maximization problem $\text{Max } f(x, y)$ such that $x^2 + y^2 - 5 \leq 0$; $x^3 + 3x^2y - 10 \leq 0$; $2x + 3y = 0$. Find out if strong duality can hold. State assumptions, if any, are required.



ChatGPT:

The maximization problem you provided is a constrained optimization problem, and it's common to use the concept of duality in such problems, particularly from the field of convex optimization. The concept of "strong duality" usually applies to convex optimization problems, where the primal and dual optimal objectives match exactly. Here's how to evaluate whether strong duality holds for the given problem.

The Problem:

Maximize $f(x, y)$ subject to:

- $x^2 + y^2 - 5 \leq 0$
- $x^3 + 3x^2y - 10 \leq 0$
- $2x + 3y = 0$

Conditions for Strong Duality:

- 1. Convexity of Objective Function:** The objective function $f(x, y)$ needs to be convex for strong duality to hold. You didn't specify the form of $f(x, y)$, so this is a necessary assumption.
- 2. Convexity of Constraints:**
 - The constraint $x^2 + y^2 - 5 \leq 0$ defines a convex region (a circle centered at the origin with a radius of $\sqrt{5}$).
 - The constraint $x^3 + 3x^2y - 10 \leq 0$ is generally not a convex function, given that the second mixed partial derivative changes sign under certain conditions.
 - The constraint $2x + 3y = 0$ is an affine function, which is both convex and concave.
- 3. Slater's Condition:** For strong duality to hold, Slater's condition states that there should exist some x, y such that all inequality constraints are strictly feasible (i.e., they are satisfied with strict inequalities). Given that the problem involves polynomial inequalities, identifying whether such an x, y exists could be complex.

Summary:

- The objective function's form isn't provided, so we assume it's convex.
- The constraints are not all convex, so typical strong duality conditions for convex optimization don't apply directly.
- Slater's condition is unclear due to the complexity of the inequalities.

Therefore, based on the given information, we cannot confirm that strong duality holds for this optimization problem. Additional assumptions about the objective function and constraints would be needed to make such a determination.



You:

Show that S and S^* have the same eigenvalues. If u is an eigenvector of S^* , can you find the eigenvector of S ? If the vector unique?



ChatGPT:

Let S be a square matrix and S^* be its conjugate transpose (also known as the Hermitian adjoint).

Eigenvalues of S and S^* :

Suppose λ is an eigenvalue of S with corresponding eigenvector v such that $Sv = \lambda v$. Taking the conjugate transpose of both sides yields:

$$(Sv)^* = (\lambda v)^* v^* S^* = \bar{\lambda} v^*$$

$$(Sv)^* = (\lambda v)^* v^* S^* = \bar{\lambda} v^*$$

Multiplying both sides by λ^{-1} :

$$\lambda v^* S^* = \bar{\lambda}^2 v^*$$

$$\lambda^{-1} v^* S^* = \bar{\lambda} v^*$$

Since $\bar{\lambda}^2 \lambda^{-2}$ is a scalar, $\bar{\lambda} v^* \lambda^{-1}$ is an eigenvector of $S^* S^*$ with the eigenvalue $\bar{\lambda}^2 \lambda^{-2}$.

Similarly, if μ is an eigenvalue of $S^* S^*$ with corresponding eigenvector u such that $S^* u = \mu u$, it can be shown that μ is an eigenvalue of S as well.

Therefore, S and S^* have the same eigenvalues, although possibly with different multiplicities.

Eigenvector of S :

If u is an eigenvector of $S^* S^*$ corresponding to eigenvalue μ , then:

$$S^* u = \mu u$$

$$S^* u = \mu u$$

Taking the conjugate transpose:

$$u^* S = \bar{\mu} u^*$$

$$u^* S = \bar{\mu} u^*$$

Since $\bar{\mu}$ is the conjugate of an eigenvalue μ of $S^* S^*$, it must also be an eigenvalue of S . The corresponding eigenvector would be u^* .

Uniqueness of Eigenvectors:

Eigenvectors corresponding to a given eigenvalue are generally not unique. Any nonzero scalar multiple of an eigenvector is also an eigenvector corresponding to the same eigenvalue. Therefore, eigenvectors are unique only up to a scalar multiple.

In summary, if u is an eigenvector of $S^* S^*$, then its conjugate transpose u^* can serve as an eigenvector of S , both corresponding to the same eigenvalue μ (or its conjugate $\bar{\mu}$). The eigenvector is unique up to a scalar multiple.



You:

Take the function to be minimized in SVM to be $\|w\|$ and derive the conditions.



ChatGPT:

In Support Vector Machines (SVM), the primal problem aims to find a hyperplane that maximizes the margin while correctly classifying the data points. The margin is the perpendicular distance from the hyperplane to the closest data points from each class.

The primal problem is usually formulated as follows:

$$\text{Minimize } \frac{1}{2} \|w\|^2$$

$$\text{Minimize } \frac{1}{2} \|w\|^2$$

$$\text{subject to } y_i(w \cdot x_i + b) - 1 \geq 0 \quad \forall i$$

$$\text{subject to } y_i(w \cdot x_i + b) - 1 \geq 0 \quad \forall i$$

Here, $\|w\|$ is the norm of the vector w , and w and b define the hyperplane. The vectors x_i are the feature vectors of the data points, and y_i are the corresponding labels which are either $+1$ or -1 .

In your case, you want to minimize $\|w\|$ instead of $\frac{1}{2} \|w\|^2$. The primal problem becomes:

$$\begin{aligned} & \text{Minimize} \quad \|w\| \\ & \text{subject to} \quad y_i(w \cdot x_i + b) - 1 \geq 0 \quad \forall i \end{aligned}$$

Deriving the Conditions:

- Objective Function:** The objective function is $\|w\|$, which is not differentiable everywhere. This could make the optimization more challenging compared to $\frac{1}{2} \|w\|^2$.
- Constraints:** The constraint $y_i(w \cdot x_i + b) - 1 \geq 0$ ensures that each data point x_i is on the correct side of the hyperplane and outside the margin. For the points on the edge of the margin, $y_i(w \cdot x_i + b) - 1 = 0$.
- Feasibility:** The problem is feasible if there exists some hyperplane that can classify all the data points according to these constraints. In the case of non-linearly separable data, one would generally introduce slack variables and move to a "soft-margin" SVM.
- Optimality Conditions:** The KKT (Karush-Kuhn-Tucker) conditions will give you the conditions for optimality in constrained optimization. However, as the objective function $\|w\|$ is not differentiable at $w = 0$, typical KKT analysis may not directly apply.

To summarize, minimizing $\|w\|$ subject to the given constraints can make the problem more difficult to solve compared to the standard SVM objective of minimizing $\frac{1}{2} \|w\|^2$, largely because the objective function is not differentiable everywhere. Nonetheless, the problem structure remains similar: you're aiming to find the hyperplane that can correctly classify the data points while minimizing the norm of w .