A Linear Algebra Review

Taken largely from a chapter written by Chung-Ming Kuan and published online in 2002

Basics

A matrix is an array of numbers

\[
A = \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

\( m \times n \) matrix with \( m \) rows and \( n \) columns

- \( a_{ij} \) is the entry in the \( i \)th row and \( j \)th column

- An \( n \times 1 \) matrix is a column vector
- An \( 1 \times n \) matrix is a row vector

For a matrix \( A \), \( a_i \) is its \( i \)th column

Definitions

- A square matrix has an equal number of rows and columns.
- A diagonal matrix is all zeros except for the diagonal elements.
- A diagonal matrix whose diagonal elements are all 1 is an identity matrix.
  Usually denoted \( I \), \( I_m \) for the \( m \)-dimensional identity
- \( 0 \) is the matrix of all 0s
If \( f \) is a vector-valued function \( f: \mathbb{R}^n \to \mathbb{R}^m \), \( \nabla_\Theta f(\Theta) \) is an \( m \times n \) matrix of the first derivatives of \( f \) w.r.t. the elements of \( \Theta \):

\[
\nabla_\Theta f(\Theta) = \begin{bmatrix}
\frac{\delta f_1(\Theta)}{\delta \Theta_1} & \frac{\delta f_1(\Theta)}{\delta \Theta_2} & \ldots & \frac{\delta f_1(\Theta)}{\delta \Theta_n} \\
\frac{\delta f_2(\Theta)}{\delta \Theta_1} & \frac{\delta f_2(\Theta)}{\delta \Theta_2} & \ldots & \frac{\delta f_2(\Theta)}{\delta \Theta_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\delta f_m(\Theta)}{\delta \Theta_1} & \frac{\delta f_m(\Theta)}{\delta \Theta_2} & \ldots & \frac{\delta f_m(\Theta)}{\delta \Theta_n}
\end{bmatrix}
\]

if \( n = 1 \) then the gradient of \( f \) is a column vector. Also, if \( n = 1 \) the \( m \times m \) Hessian matrix is

\[
\nabla_\Theta^2 f(\Theta) = \nabla_\Theta (\nabla_\Theta f(\Theta)) = \begin{bmatrix}
\frac{\delta^2 f(\Theta)}{\delta \Theta_1^2} & \frac{\delta^2 f(\Theta)}{\delta \Theta_1 \delta \Theta_2} & \ldots & \frac{\delta^2 f(\Theta)}{\delta \Theta_1 \delta \Theta_n} \\
\frac{\delta^2 f(\Theta)}{\delta \Theta_2 \delta \Theta_1} & \frac{\delta^2 f(\Theta)}{\delta \Theta_2^2} & \ldots & \frac{\delta^2 f(\Theta)}{\delta \Theta_2 \delta \Theta_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\delta^2 f(\Theta)}{\delta \Theta_n \delta \Theta_1} & \frac{\delta^2 f(\Theta)}{\delta \Theta_n \delta \Theta_2} & \ldots & \frac{\delta^2 f(\Theta)}{\delta \Theta_n^2}
\end{bmatrix}
\]

Two matrices are of equal size if they have the same number of rows and columns. Matrix equality is defined for matrices of the same size. (obvious)

The transpose of a matrix \( A \) is denoted \( A^T \) or \( A' \)

the \( ij \)th element of \( A^T \) is the \( ji \)th element of \( A \)

A matrix whose transpose is equal to itself is symmetric.

Two matrices of the same size can be added:

\[ C = A + B = B + A \]

Obviously, \( A + O = A \) and \( (A + B) + C = A + (B + C) \)
\[(AB)' = B' A'\]

Proof:

\[c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}\]

\[c_{ij} = \sum_{k=1}^{n} b'_{ik} a'_{kj}\]

\[\begin{bmatrix} C \end{bmatrix} = \sum_{k=1}^{n} \begin{bmatrix} B' \end{bmatrix} \begin{bmatrix} A' \end{bmatrix}\]

\[\begin{bmatrix} C \end{bmatrix} = \sum_{k=1}^{n} \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} B \end{bmatrix}\]

\[b'_{ik} = b_{kj}, \quad a'_{kj} = a_{ik}\]
A scalar \( c \) times \( A \) is written:

\[
cA = \{ c a_{ij} \}
\]

Obviously \( cA = Ac \) and \(-1A = -A \) and \( A A = A \)

**Matrix multiplication**

\( AB \) is only defined if the number of columns of \( A \) is the same as the number of rows of \( B \)

\[
m \times n \begin{bmatrix} A \\ \end{bmatrix} \times n \begin{bmatrix} B \\ \end{bmatrix} = n \begin{bmatrix} C \\ \end{bmatrix}
\]

where

\[
c_{mp} = \sum_{i=1}^{n} a_{mi} \cdot b_{ip}
\]

Obviously \( AB \neq BA \) (in general), however

\[
A(BC) = (AB)C \quad \text{associative}
\]

\[
A(B+C) = AB + AC \quad \text{commutative}
\]

One can verify \((AB)' = B'A'\)

**Inner product of vectors**

If \( \vec{x} \) and \( \vec{z} \) are vectors their inner product is:

\[
\vec{x} \cdot \vec{z} = \sum_{i=1}^{n} x_i z_i
\]

If \( \vec{x} \) is \( m \)-di- and \( \vec{z} \) is \( n \)-di-

the outer product of \( \vec{x} \) and \( \vec{z} \) is

\[
\vec{x} \vec{z}^T \quad \text{is a matrix with elements}
\]

\[
\{ y_{ij} \} \quad 1 \leq i \leq m, 1 \leq j \leq n
\]
The Euclidean norm of a vector \( \mathbf{v} \) (its "length") is
\[
\| \mathbf{v} \| = \left( \sum_{i=1}^{n} v_i^2 \right)^{1/2} = \left( \frac{\mathbf{v} \cdot \mathbf{v}}{2} \right)^{1/2}
\]

\( \mathbf{0} \) has Euclidean norm 0, a unit vector has Euclidean norm 1. Examples?

Orthogonality:
\[
\begin{align*}
\mathbf{y} \\
\mathbf{z} \\
\mathbf{y} \cdot \mathbf{z} = 0
\end{align*}
\]

Law of Cosine
\[
\| \mathbf{y} - \mathbf{z} \|^2 = \| \mathbf{y} \|^2 + \| \mathbf{z} \|^2 - 2 \| \mathbf{y} \| \| \mathbf{z} \| \cos \theta
\]

\[
\left( \mathbf{y} - \mathbf{z} \right) \cdot \left( \mathbf{y} - \mathbf{z} \right) = \| \mathbf{y} \|^2 + \| \mathbf{z} \|^2 - 2 \| \mathbf{y} \| \| \mathbf{z} \| \cos \theta
\]

\[
\Rightarrow \quad \| \mathbf{y} \|^2 - \| \mathbf{z} \|^2 = \| \mathbf{y} \| \| \mathbf{z} \| \cos \theta
\]

which says that is \( \theta = \frac{\pi}{2} \) (90°) then \( \mathbf{y} \cdot \mathbf{z} = 0 \).

In this case \( \mathbf{y} \) and \( \mathbf{z} \) are said to be orthogonal.

A square matrix \( A \) is orthogonal if
\[
A' A = A A' = I
\]

i.e. each row (and column) of \( A \) is a unit vector and orthogonal to the others.

If \( \mathbf{y} = c \mathbf{z} \) for some \( c \neq 0 \)
\( \mathbf{y} \) and \( \mathbf{z} \) are said to be linearly dependent.
Consider differentiation w.r.t. vectors and matrices

Let \( a, a' \) be two \( D \)-dim. vectors, then

\[
\nabla_\theta (a') = a
\]

Check:

\[
\nabla_\theta (a') = \left[ \frac{d a'}{d \theta_1}, \frac{d a'}{d \theta_2}, \ldots, \frac{d a'}{d \theta_D} \right]
\]

\[
= \left[ a_1, a_2, \ldots, a_D \right] = a
\]

\[
\nabla_\theta (\theta' \theta) = 2 \theta'A \theta \quad \text{where } A \text{ is a } D \times D \text{ matrix}
\]

Check:

\[
\nabla_\theta (\theta' \theta) = \left[ \frac{d \theta' \theta}{d \theta_1}, \frac{d \theta' \theta}{d \theta_2}, \ldots, \frac{d \theta' \theta}{d \theta_D} \right]
\]

Note \( \theta' \theta \) looks like

\[
\begin{bmatrix}
    \theta_1, & \ldots, & \theta_D
\end{bmatrix}
\begin{bmatrix}
    a_{11} & a_{12} & \ldots & a_{1D} \\
    a_{21} & a_{22} & \ldots & a_{2D} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{D1} & a_{D2} & \ldots & a_{DD}
\end{bmatrix}
\begin{bmatrix}
    \theta_1 \\
    \theta_2 \\
    \vdots \\
    \theta_D
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    \theta_1 a_{11} + \theta_2 a_{12} + \ldots + \theta_D a_{1D} \\
    \theta_1 a_{21} + \theta_2 a_{22} + \ldots + \theta_D a_{2D} \\
    \vdots \\
    \theta_1 a_{D1} + \theta_2 a_{D2} + \ldots + \theta_D a_{DD}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    \theta_1 a_{11} + \theta_2 a_{12} + \ldots + \theta_D a_{1D} \\
    \theta_1 a_{21} + \theta_2 a_{22} + \ldots + \theta_D a_{2D} \\
    \vdots \\
    \theta_1 a_{D1} + \theta_2 a_{D2} + \ldots + \theta_D a_{DD}
\end{bmatrix}
\]

\[
\theta_1 a_{11} + \ldots + \theta_D a_{DD}
\]
Taking the derivative of this with to \( \Theta \) repeatedly yields

\[
\frac{\partial \Theta^\prime \Theta}{\partial \Theta} = \begin{bmatrix} 0, \cdots, -2 \Theta a_{ii}, \cdots, 0 \end{bmatrix}
\]

So

\[
\nabla \Theta^\prime \Theta = 2 \Theta A
\]

Combining these two rules yields

\[
\nabla^2 \Theta^\prime A \Theta = 2 A
\]

---

**Matrix Determinant**

Let \( A \) be a square matrix, and let \( A_{ij} \) be the submatrix of \( A \) after deleting row \( i \) and column \( j \). The determinant of \( A \) is

\[
|A| = \text{det}(A) = \sum_{i+j=1}^{m} (-1)^{i+j} a_{ij} \text{det}(A_{ij})
\]

or equivalently

base case \( |A| = a_{11} a_{22} - a_{12} a_{21} \)

\[
|A| = \text{det}(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \text{det}(A_{ij})
\]

A square matrix with a non-zero determinant is called **nonsingular**; otherwise it is called **singular**.

Fact: \( \text{det}(A) = \text{det}(A^\prime) \)

Sketch proof: base case is invariant to transpose, since \( \text{det} \) is equivalent to both row and column expansion, \( 3 \times 3 \) det. is clearly invariant to transpose. Induction from there.
Another super useful fact:

$$|cA| = c^n |A|$$

Proof: again, base case is $2 \times 2$ 

$$|A| = c_{11}c_{22} - c_{12}c_{21} = c^2 |A|$$

from definition $3 \times 3$ will be $c^3 |A|$, induction.

Without proof:

$$|AB| = |A||B| = |BA|$$

Many facts at once: if $A$ is orthogonal
the $AA' = I$.

From definition of det it is easy to see that $|I| = 1$, so

$$|I| = |AA'| = |(A^T)|^2 \Rightarrow |A| = 1 \text{ or } -1$$

i.e. orthogonal matrices must have det $1$ or $-1$.

Trace

The trace of a matrix is the sum of its diagonal elements.

$$\text{trace}(A) = \sum_{i=1}^{n} a_{ii}$$

$$\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B)$$

$$\text{trace}(cA + dB) = c \text{ trace}(A) + d \text{ trace}(B)$$

Without proof:

$$\text{trace}(AB) = \text{trace}(BA)$$

if both products are well defined.
Matrix Inverses

A non-singular matrix $A$ has a unique inverse $A^{-1}$ such that $AA^{-1} = A^{-1}A = I$.

Matrix inversion and transposition can be interchanged, i.e.,

$$(A')^{-1} = (A^{-1})'$$

Because

$$ABB^{-1}A^{-1} = I$$

with $A$ and $B$ non-singular and compatible,

we have

$$(AB)^{-1} = B^{-1}A^{-1}$$

Most matrix inverses must be computed but some are easy. If $A$ is diagonal then $A^{-1}$ is diagonal with diagonal elements that are the reciprocal of the original. If $A$ is orthogonal then because $AA' = I$ and $A'A = I$.

Partitioned matrices can be inverted easily sometimes.

Matrix Rank (important in lin. alg.)

A set of vectors $\vec{z}_1, \ldots, \vec{z}_n$ are linearly independent if $c_1, c_2, \ldots, c_n = 0$ is the only solution to

$$c_1\vec{z}_1 + c_2\vec{z}_2 + \cdots + c_n\vec{z}_n = 0$$

otherwise they are linearly dependent. When 2 vectors are linearly dependent they lie on the same line; 3, plane or line; 4, plane, line, or volume.
Claim: The column rank and row rank of a matrix are equal.

Pf: The column rank of a matrix \( A \) is the max. number of linearly independent column vectors of \( A \) (row rank def. the same). If the col. or row. rank of \( A \) equals the corresponding dimensionality then \( A \) is said to be of full column or row rank.

The space spanned by a set of vectors \( \mathbf{z}_1, \ldots, \mathbf{z}_m \) is the collection of all linear combinations of those vectors and is denoted \( \text{span}\{\mathbf{z}_1, \ldots, \mathbf{z}_m\} \).

The space spanned by the column vectors of \( A \) is \( \text{span}(A) \) and is called the column space of \( A \). \( \text{span}(A^\top) \) is the row space of \( A \).

Let \( A \) be an \( n \times k \) matrix with \( k \leq n \) and suppose \( r = \text{row rank}(A) \leq n \) and \( c = \text{column rank}(A) \leq k \).

Assume w.l.o.g. the first \( r \) rows of \( A \) are lin. independent then all rows of \( A \) can be expressed as

\[
\mathbf{a}_i = q_{i1} \mathbf{a}_1 + q_{i2} \mathbf{a}_2 + \cdots + q_{ir} \mathbf{a}_r
\]

where the \( j \)th element of \( \mathbf{a}_i \) is

\[
q_{ij} = q_{i1} a_{1j} + q_{i2} a_{2j} + \cdots + q_{ir} a_{rj}
\]

but from this it is clear that any column vector of \( A \) can be written as the linear combination of the \( r \) vectors

\[
\mathbf{a}_j = \mathbf{c}_1 c_1 + \mathbf{c}_2 c_2 + \cdots + \mathbf{c}_r c_r
\]
This uses the column rank of $A$ must also be less than or equal to $r$. The same arg. can be applied to the transpose of $A$, or, starting with the column vectors yielding the result that the row rank of $A$ must be less or equal to $c$.

$$\Rightarrow \text{row rank}(A) = \text{column rank}(A) = \text{rank}(A)$$

Clearly: $\text{rank}(A) = \text{rank}(A')$

if $\text{rank}(A) = n$ and $A$ is $n \times n$ then $A$ is full rank.

Fact: A full rank matrix is nonsingular and vice versa.

It can be shown that:

$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$$

$$\text{rank}(AB) \leq \min \left[ \text{rank}(A), \text{rank}(B) \right]$$

Using these we have, when $A$ is non-singular:

$$\text{rank}(AB) \leq \text{rank}(B) = \text{rank}(A'AB) \leq \text{rank}(AB)$$

$$\Rightarrow \text{rank}(AB) = \text{rank}(B) \text{ when } A \text{ non-singular}$$

the other direction works as well

$$\text{rank}(BC) = \text{rank}(B) \text{ when } C \text{ non-singular}$$
Eigenvalues - Eigen vectors

Given a square matrix \( A \) if

\[ A \vec{z} = \lambda \vec{z} \]

for some scalar \( \lambda \) and non-zero vector \( \vec{z} \)

then

\( \vec{z} \) is an eigenvector of \( A \) corresponding to \( \lambda \) an eigenvalue

Eigen vectors and eigenvalues are particularly interpretable when \( A \) is a rotation or reflection matrix, the eigen vectors then are the axes of rotation and the eigen values sign indicate reflection through some space.

Given an eigenvalue \( \lambda \), let \( \vec{z}_1, \ldots, \vec{z}_k \) be associated eigen vectors. Then,

\[ A(a_1 \vec{z}_1 + a_2 \vec{z}_2 + \cdots + a_k \vec{z}_k) = \lambda(a_1 \vec{z}_1 + a_2 \vec{z}_2 + \cdots + a_k \vec{z}_k) \]

so any linear combination of the \( \vec{z} \)'s is again an eigen vector and the set of all such vectors is an eigen space associated with eigenvalue \( \lambda \).

If \( A \) (\( nxn \)) has distinct eigen values, each eigenvalue must correspond to one eigen vector and the set of such eigen vectors must be linearly independent.

Since the choice of eigenvector is identifiable only up to a constant, often unit length eigenvectors are chosen.
\[ \mathbb{E} \left[ (x - \mu)^T (x - \mu) \right] \]
Let \( C \) be the matrix consisting of three \( n \), distinct, unit length, linearly independent eigen vectors. Clearly \( C \) is non-singular, that means we can write

\[
AC = \lambda C
\]

where

\[
\lambda = \begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\vdots \\
\lambda_n
\end{bmatrix}
\]

and using the non-singularity of \( C \) we have

\[
C^\top AC = \lambda \quad \text{or} \quad A = C \Lambda C^{-1}
\]

Another, when \( A \) has \( n \) distinct eigenvalues

\[
\det(A) = \det(C \Lambda C^{-1}) = \det(\Lambda) \det(C) \det(C^{-1}) = \det(\Lambda) = \prod_{i=1}^{n} \lambda_i
\]

\[
\text{trace}(A) = \text{trace}(C \Lambda C^{-1}) = \text{trace}(C^\top C \Lambda) = \text{trace}(\Lambda) = \sum_{i=1}^{n} \lambda_i
\]

When \( A = C \Lambda C^{-1} \) we have \( A^{-1} = C \Lambda^{-1} C^{-1} \)

so the eigen vectors of \( A^{-1} \) are the same as those of \( A \), the eigen values of \( A^{-1} \) are the reciprocals of the eigen values of \( A \).
Symmetric Matrices

Let \( z_1 \) and \( z_2 \) be two eigenvectors of \( A \) corresponding to distinct eigenvalues \( \lambda_1 \neq \lambda_2 \).

If \( A \) is symmetric,

\[
z_2^T A z_1 = \lambda_1 z_2^T z_1 = \lambda_2 z_2^T z_1,
\]

which because \( \lambda_1 \neq \lambda_2 \) \( \Rightarrow \) \( z_2^T z_1 = 0 \).

A symmetric matrix is orthogonally diagonalizable such that

\[
C^T AC = \Lambda \quad \text{or} \quad A = C \Lambda C^T
\]

where \( \Lambda \) is a diagonal matrix of eigenvalues and \( C \) is the orthogonal matrix of associated eigenvectors.

Remember: non-singular transforms preserve rank.

So, for a symmetric matrix \( A \), \( \text{rank}(C \Lambda) = \text{rank}(A) \), i.e. the number of non-zero eigenvalues of \( A \).
Lemma: Let $A$ be an $n \times n$ symmetric matrix. Then

$$\det(A) = \det(L) = \prod_{i=1}^{n} \lambda_i,$$

$$\text{trace}(A) = \text{trace}(L) = \sum_{i=1}^{n} \lambda_i.$$

Obviously, a symmetric matrix is non-singular if all its eigenvalues are non-zero.

A symmetric matrix $A$ is said to be positive definite if $b'Ab > 0 \quad \forall \ b \neq 0$. Pos. semi-def. if $b'Ab \geq 0$.

If $A$ is pos. def. $\Rightarrow$ $A$ non-singular.
A pos. semi-def. $\Rightarrow$ """"", $A$ may be singular.

If $A$ is symmetric and orthogonally diagonalized as $C'AC = \Lambda$ and if $A$ is p.s.d. then for $\tilde{b} \neq 0$

$$\tilde{b}' \Lambda \tilde{b} = \tilde{b}'(C'AC)\tilde{b} = \tilde{b}'A\tilde{b} \geq 0$$

where $\tilde{b} = Cb$. This shows that $\Lambda$ is also p.s.d. and $\Lambda$'s diagonal elements must be positive.

Lemma: A symmetric matrix is p(s)d iff. its eigenvalues are all positive (non-negative).
For a symmetric positive definite matrix $A$, $A^{-1/2}$ is such that $A^{-1/2}A^{-1/2} = A^{-1}$. This can be arrived at via

$$A^{-1} = C \Lambda^{-1} C' = (C \Lambda^{-1/2} C')(C \Lambda^{-1/2} C')$$

so we may choose $A^{-1/2} = C \Lambda^{-1/2} C'$

**Orthogonal Projection**

A matrix $A$ is idempotent if $A^2 = A$. Given a vector $\mathbf{y}$ in Euclidean space $V$, a projection of $\mathbf{y}$ onto a subspace $S$ of $V$ is a linear transformation of $\mathbf{y}$ to $S$. The projection can be written $P\mathbf{y}$, where $P$ is a transformation matrix. Projecting a projection should not affect the projection, i.e.,

$$P(P\mathbf{y}) = P^2\mathbf{y} = P\mathbf{y}$$

The matrix $P$ is called a projection matrix if it is idempotent.

A projection of $\mathbf{y}$ onto $S$ is orthogonal if $P\mathbf{y}$ is orthogonal to the difference between $\mathbf{y}$ and $P\mathbf{y}$.

Algebraically

$$(\mathbf{y} - P\mathbf{y})' P\mathbf{y} = ((I - P)\mathbf{y})' P\mathbf{y} = (\mathbf{y}' (I - P)^2 \mathbf{y}$$

This can only be zero if $(I - P)^2 = 0$. This can only happen if $P = PP$. This shows that $P$ must be symmetric.

**Conclusion:** $P$ is an orthogonal projection matrix if $P$ is symmetric and idempotent.
If $P$ is an orthogonal projection matrix, it can clearly be seen that $I-P$ is idempotent:

$$(I-P)(I-P) = I - 2P + PP = I - P$$

Since $I-P$ is symmetric, it is also an orthogonal projection matrix.

Since $(I-P)P = 0$, the projections $Py$ and $(I-P)y$ must be orthogonal.

So, any vector $y$ can be uniquely decomposed into two orthogonal components:

$$y = Py + (I-P)y$$

If $A$ is symmetric and idempotent and $C$ is orthogonal, then

$$\Lambda = C'AC = C'ACC'AC = \Lambda^2$$

This can only happen if the entries of $\Lambda$ are $0$ or $1$. 
This implies since the rank of $A'A$ and $AA'$ and $A$ are all the same, i.e.
\[ \text{rank}(A) = \text{rank}(A'A) = \text{rank}(AA') \]

If $A$ is full column rank $\Rightarrow \text{rank}(A'A) = k$ w/ $k$ pos. eigs.
Facts: A symmetric and idempotent matrix is positive semi-definite with eigenvalues 0 or 1.
- Trace (A) is the number of non-zero eigenvalues of A and hence rank (A) = trace (A).

Remember: if A is symmetric rank (A) = rank (A') and trace (A) = trace (A').

Combining these two yields:

Fact: For a symmetric and idempotent matrix A, rank (A) = trace (A) the number of non-zero eigenvalues of A.

Let A be an n x k matrix. Clearly A'A and AA' are symmetric.

If A is full column rank k < n, 
P = A(A'A)'A' is symmetric and idempotent and as such an orthogonal projection matrix.

As trace (P) = trace (A'A (A'A)') = trace (I_k) = k
we see that P as k eigenvalues equal to 1 and as such rank (P) = k. Similarly, rank (I-P) = n-k.

---

*homework*