



FROM QUADRATIC FORM TO SINGULAR VALUE DECOMPOSITION

TUTORIAL-STYLE NOTES I WROTE FOR MYSELF
BASED ON LAY'S "LINEAR ALGEBRA AND ITS APPLICATIONS"

WRITTEN BY

LEO MAVROPALIAS

COMPILED WITH

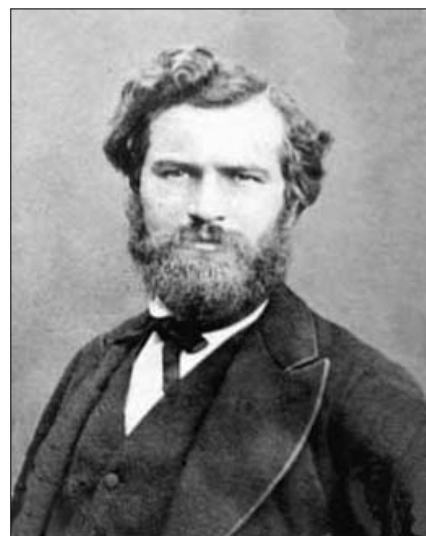
`overleaf.com`



J. Sylvester
1814 – 1897



E. Beltrami
1835 - 1900



C. Jordan
1838 – 1922

DRAFT VERSION 0.935
SEPTEMBER 10, 2018



Contents

1	Quadratic form	2
1.1	From matrix to quadratic form	2
1.2	From quadratic form to matrix	3
1.3	Change of variable in quadratic form	4
2	Principal axes theorem	5
2.1	A geometric view of the principal axes theorem	5
3	Classifying quadratic surfaces	8
4	Constrained optimisation	13
4.1	Optimisation of quadratic form in the unit sphere	13
4.2	(Optional) Optimisation of general 2nd order form	16
4.2.1	Unbounded quadric problem (part 1)	16
4.2.2	Scalar quadratic optimization without constraints	17
4.2.3	Unbounded quadratic problem (part 2)	17
4.2.4	Unconstrained optimisation using analytic methods; the Hessian matrix	18
4.2.5	Constrained optimisation with linear constraint	20
4.2.6	General constrained optimisation	22
5	Singular value decomposition (SVD)	22
5.1	Introduction	22
5.1.1	Geometric interpretation of SVD	27
5.2	Application 1: SVD and fundamental subspaces	30
5.3	Application 2: Low rank approximation	31
5.3.1	SVD and norms	32
5.4	Application 3: Image compression	33
5.4.1	Application 3.1: Image deblurring	34
5.5	Application 4: Pseudoinverse and least squares	34
5.6	Application 5: Principal Component Analysis (PCA)	35
5.6.1	Application of PCA for facial recognition	35
A	Appendices	36
A.1	Appendix 1. Matrix form of general conic section equation	37
A.2	Appendix 2. 3D paraboloid equations	38
A.3	Appendix 3. Cholesky factorisation	39
A.4	Properties of the gradient operator	41
A.5	Taylor series for multi variable functions	43
A.6	Intuition behind Lagrange multiplier	46
A.7	Relationship between rank and number of non-zero eigenvalues	47
A.8	Why do orthogonal matrices represent rotations?	48
A.9	Matlab program to compress image by using SVD	49
A.10	Linear least squares solution	50
A.11	Pseudoinverse with QR factorisation	53

1 Quadratic form

DEFINITION 1.1. A quadratic form is a polynomial in several variables where each term is of total degree two. In other words it is linear combination of products of two variables (possibly repeated variables).

A quadratic form can neatly be written as a product of a matrix and two vectors. Example of quadratic forms are the following:

$$\begin{aligned} Q(x_1, x_2) &= -x_1^2 + 3x_2^2, \\ Q(x_1, x_2, x_3) &= 5x_1^2 - 4x_3^2 - x_1x_2 + 7x_2x_3 \\ Q(x_1, x_2, x_3) &= -x_1x_2 + 7x_2x_3 + 2x_1x_3 \end{aligned}$$

A quadratic form can neatly be written as a product of a matrix that describes the polynomial coefficients with two vectors that store the free variables. For example, take the product

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2, & 2x_1 + 4x_2 \end{bmatrix}$$

Now multiply with the column vector $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$.

$$\begin{bmatrix} x_1 + 3x_2, & 2x_1 + 4x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 5x_1x_2 + 4x_2^2$$

Note that the off-diagonal entries themselves do not affect the result, but their sum does. For example, $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, $\begin{bmatrix} 1 & 2.5 \\ 2.5 & 4 \end{bmatrix}$ would all product the same polynomial. When we want to find which matrix produced the form, we will prefer the symmetric case.

COROLLARY 1.1. A quadratic form can be written as a function $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ in the form

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \tag{1.1}$$

, where \mathbf{A} an $n \times n$ real symmetric matrix.

As mentioned before, \mathbf{x} is a column vector. Below we will discuss how we can quickly go from matrix \mathbf{A} to its quadratic form polynomial $Q(\mathbf{x})$ and vice versa. We assume, for ease, that \mathbf{A} is 2×2 .

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

, with $b = c$ because \mathbf{A} is symmetric.

1.1 From matrix to quadratic form

Let's find $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ given $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} ax_1 + cx_2 & bx_1 + dx_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= ax_1^2 + (b + c)x_1x_2 + x_2^2 \end{aligned}$$

Therefore quadratic form is a linear combination of the squares x_1^2, x_2^2 and the cross-product term (terms in general) x_1x_2 . Notice how for matrix \mathbf{A} the diagonal elements ($i = j$) scale the squares and the off-diagonal ones ($i \neq j$) scale the cross-product term.

LEMMA 1.1. Matrices \mathbf{A} and $\hat{\mathbf{A}} = \frac{\mathbf{A} + \mathbf{A}^T}{2}$ have the same quadratic form.

LEMMA 1.2. Matrices \mathbf{A} and \mathbf{A}^T have the same quadratic form, i.e.

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{A}^T \mathbf{x} \quad (1.2)$$

Proof. It is left to the reader. Verify it first for a 2×2 matrix. □

Proof. For matrix $\hat{\mathbf{A}}$:

$$\begin{aligned} Q(\mathbf{x}) &= \mathbf{x}^T \frac{\mathbf{A} + \mathbf{A}^T}{2} \mathbf{x} \\ &= \frac{1}{2} (\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{A}^T \mathbf{x}) \\ &= \frac{1}{2} (\mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{A} \mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \end{aligned}$$

□

EXAMPLE 1.1. If $\mathbf{A} = \begin{bmatrix} 3 & -1 \\ -3 & 7 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 7 \\ -2 & 7 & 3 \end{bmatrix}$

find the quadratic form of

- | | |
|--------------------|---|
| (a) \mathbf{A} | (c) $\frac{\mathbf{A} + \mathbf{A}^T}{2}$ |
| (b) \mathbf{A}^T | (d) \mathbf{B} |
| | (e) \mathbf{C} |

Solution:

- (a) $Q(\mathbf{x}) = 3x_1^2 + (-1 - 3)x_1x_2 + 7x_2^2 = 3x_1^2 - 4x_1x_2 + 7x_2^2$
 (b) $Q(\mathbf{x}) = 3x_1^2 + (-3 - 1)x_1x_2 + 7x_2^2 = 3x_1^2 - 4x_1x_2 + 7x_2^2$
 (c) $\frac{\mathbf{A} + \mathbf{A}^T}{2} = \frac{1}{2} \left(\begin{bmatrix} 3 & -1 \\ -3 & 7 \end{bmatrix} + \begin{bmatrix} 3 & -3 \\ -1 & 7 \end{bmatrix} \right) = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$
 $\therefore Q(\mathbf{x}) = 3x_1^2 + (-2 - 2)x_1x_2 + 7x_2^2 = 3x_1^2 - 4x_1x_2 + 7x_2^2$
 (d) $Q(\mathbf{x}) = 3x_1^2 + 7x_2^2$
 (e) $Q(\mathbf{x}) = x_1^2 + 2x_1^2 + 3x_3^2 - 4x_1x_2 + 14x_2x_3$

■

1.2 From quadratic form to matrix

Suppose we are given that $Q(\mathbf{x}) = ax_1^2 + ex_1x_2 + dx_2^2$ for some 2×2 symmetric matrix \mathbf{A} . We saw in the previous section that the symmetric matrix that yielded that product is:

$$\mathbf{A} = \begin{bmatrix} a & \frac{e}{2} \\ \frac{e}{2} & d \end{bmatrix}$$

In general, the coefficient of the cross-product $x_i x_j$ goes in entry (i, j) . When we want to recover matrix \mathbf{A} from $Q(\mathbf{x})$, we want to recover the symmetric one and we'll see why.

EXAMPLE 1.2. If $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$ is the quadratic form of some symmetric matrix \mathbf{A} , find that matrix.

Solution:

In the main diagonal we have 5, 3, 2, For the cross-product entries we have $a_{12} = a_{21} = -1/2$, $a_{23} = a_{32} = 4$, therefore:

$$\mathbf{A} = \begin{bmatrix} 5 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}$$

■

1.3 Change of variable in quadratic form

DEFINITION 1.2. If $x \in \mathbb{R}^n$, then a change of variable is an equation of the form

$$\mathbf{x} = \mathbf{P}\mathbf{y} \iff \mathbf{y} = \mathbf{P}^{-1}\mathbf{x} \quad (1.3)$$

, where \mathbf{P} is an invertible matrix and $\mathbf{y} \in \mathbb{R}^n$.

If we rewrite the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, with the change of variable in Eq. (1.3), then we have

$$Q(\mathbf{x}) = \mathbf{y}^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{y} \quad (1.4)$$

❓ *Why is this useful?*

\mathbf{A} is symmetric so if we choose the correct matrix \mathbf{P} we can diagonalise the product $\mathbf{P}^T \mathbf{A} \mathbf{P}$, i.e. $\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{D}$, therefore we end up with a simpler quadratic form with no cross-product terms, only squares.

$$Q(\mathbf{x}) = \mathbf{y}^T \mathbf{D} \mathbf{y} \quad (1.5)$$

❓ *Which matrix should we choose to make the product diagonal?*

We know that a symmetric matrix is diagonalised by its eigenvector matrix. Therefore \mathbf{P} must contain the eigenvector as columns.

EXAMPLE 1.3. Make a change of variable to transform the quadratic form $Q(\mathbf{x}) = x_1^2 - 8x_1x_2 + 5x_2^2$ into a quadratic form with no cross-product terms [1].

Solution:

Write $Q(\mathbf{x})$ in matrix form:

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}, \quad \mathbf{A} = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

Diagonalise \mathbf{A} , so begin by finding the eigenvalues from:

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= 0 \implies \\ \det \left(\begin{bmatrix} 1-\lambda & -4 \\ -4 & -5-\lambda \end{bmatrix} \right) &= 0 \implies \\ (1-\lambda)(-5-\lambda) + 16 &= 0 \implies \\ \lambda &= 3, \quad \lambda = -7 \end{aligned}$$

Find the eigenvector for $\lambda = 3$.

$$\begin{aligned} \mathbf{A} \mathbf{u} &= \lambda \mathbf{u} \implies \\ \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= 3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \\ \begin{bmatrix} x_1 - 4x_2 \\ -4x_1 - 5x_2 \end{bmatrix} &= 3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \\ \begin{cases} x_1 - 4x_2 = 3x_1 \\ -4x_1 - 5x_2 = 3x_2 \end{cases} &\implies \\ x_1 &= -2, \quad x_2 = 1 \end{aligned}$$

, i.e. $\mathbf{u} = [1 \ -2]^T$, or normalised $\hat{\mathbf{u}} = [1/\sqrt{5} \ -2\sqrt{5}]^T$. Similarly we find the eigenvector for $\lambda = -7$ and we end up with the pairs:

$$\lambda = 3 : \hat{\mathbf{u}} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \lambda = -7 : \hat{\mathbf{u}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Notice how the eigenvectors are *perpendicular* as matrix \mathbf{A} is *real and symmetric*. The diagonalised result contains the eigenvalues in decreasing order and the matrix \mathbf{P} that diagonalises \mathbf{A} contains the corresponding

eigenvectors.

$$\mathbf{P} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

$$\therefore \mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{D} \implies$$

$$\mathbf{A} = \mathbf{P} \mathbf{D} \mathbf{P}^T$$

Perform change of variable $\mathbf{x} = \mathbf{P}\mathbf{y}$ in the quadratic product:

$$\begin{aligned} Q(\mathbf{x}) &= \mathbf{x}^T \mathbf{A} \mathbf{x} = x_1^2 - 8x_1x_2 - 5x_2^2 \stackrel{x = \mathbf{P}\mathbf{y}}{=} \\ &= \mathbf{y}^T \mathbf{D} \mathbf{y} = 3y_1^2 - 7y_2^2 \end{aligned}$$

■

❓ *So what is the process to diagonalise the quadratic form?*

1. Determine matrix \mathbf{A} from the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$
2. Calculate the orthogonal diagonalisation of \mathbf{A} (eigenvalue and eigenvector matrices \mathbf{D}, \mathbf{P}). Normalise the eigenvectors.
3. Perform change of variable $\mathbf{x} = \mathbf{P}\mathbf{y}$ in the quadratic form to obtain $\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{D} \mathbf{y}$.

2 Principal axes theorem

In section 1, we concluded the following.

THEOREM 2.1. *Let \mathbf{A} be an $n \times n$ real symmetric matrix. Then there exists an orthogonal change of variable $\mathbf{x} = \mathbf{P}\mathbf{y}$ that transforms the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ into a quadratic form $\mathbf{y}^T \mathbf{A} \mathbf{y}$ with no cross-product term.*

DEFINITION 2.1. *The columns of \mathbf{P} are called principal axes of the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$.*

The vector \mathbf{y} is the coordinate vector of \mathbf{x} relative to the orthonormal basis of \mathbb{R}^n given by these principal axes.

2.1 A geometric view of the principal axes theorem

If \mathbf{A} is 2×2 and $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ is its quadratic form, then if we expand it we can see that all points that satisfy

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = c \tag{2.1}$$

correspond either to

- an ellipse
- a hyperbola
- a parabola
- their degenerate cases, i.e. two intersecting lines or a single point.

Assuming $AC \neq 0$, the equation of an ellipse or hyperbola centred at the origin with axes x', y' (*standard form*) is

$$\begin{aligned} Ax^2 + Cy^2 &= 1 \\ \mathbf{x}^T \begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix} \mathbf{x} &= 1 \end{aligned} \tag{2.2}$$

The equation of an ellipse or hyperbola rotated w.r.t. axes x', y' (*non-standard form*) is

$$\begin{aligned} Ax^2 + Bxy + Cy^2 &= 1 \iff \\ \mathbf{x}^T \begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix} \mathbf{x} &= 1 \end{aligned} \tag{2.3}$$

Constants A, B, C determine the type of conic (Appendix 1), if it is not degenerate, particularly if we define our characteristic matrix $\mathbf{A} = \begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}$:

- If $B^2 - 4AC > 0 \iff \det(\mathbf{A}) < 0$, then A, C have the same sign and it is an ellipse.
- If $B^2 - 4AC < 0 \iff \det(\mathbf{A}) > 0$, then A, C have opposite signs and it is a hyperbola.
- If $B^2 - 4AC = 0 \iff \det(\mathbf{A}) = 0$, then either $A = 0$ or $C = 0$ have therefore the conic is a parabola.

Therefore the presence of the term Bxy determines whether the locus is rotated w.r.t the standard axes or not. Why that is the case is explained in Appendix 1. Some examples illustrating what is mentioned are shown below.

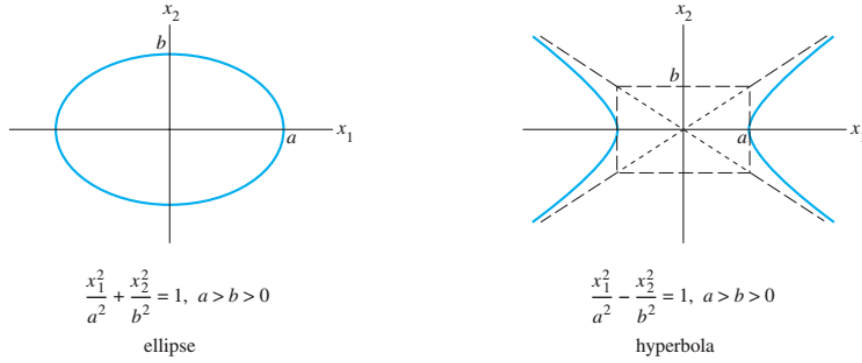


Figure 4: An ellipse and a hyperbola in standard position [1]

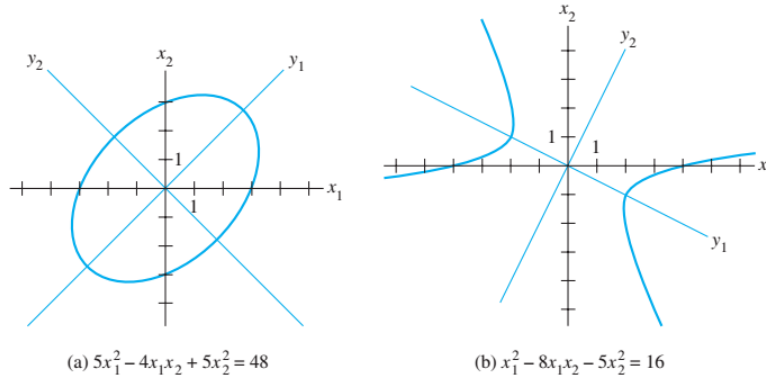
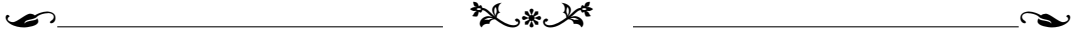


Figure 5: An ellipse and a hyperbola in non-standard (rotated) position [1]



Assuming that the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ is an ellipse, we will show that making the substitution $\mathbf{y} = \mathbf{P} \mathbf{x}$ has the effect of finding some new axes y_1, y_2 , with respect to which the conic is in standard position. We will also express the length of the ellipse axes w.r.t. the eigenvalues of A .

COROLLARY 2.1. *The primary and secondary axes point to the direction of the largest and smallest eigenvector respectively.*

COROLLARY 2.2. *The length of the primary and secondary axes of the standard ellipse given by the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is $\sqrt{\lambda_1}, \sqrt{\lambda_2}$ respectively, where $\lambda_1 \leq \lambda_2$.*

Proof. [2] We write matrix \mathbf{A} in its diagonalised form:

$$A = [\mathbf{u}_1 \quad \mathbf{u}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{bmatrix}$$

Then its quadratic form becomes:

$$\begin{aligned} Q(\mathbf{x}) &= \mathbf{x}^T \mathbf{A} \mathbf{x} \\ &= \mathbf{x}^T \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \end{bmatrix} \mathbf{x} \\ &= \lambda_1 \mathbf{x}^T \mathbf{u}_1 \mathbf{u}_1^T \mathbf{x} + \lambda_2 \mathbf{x}^T \mathbf{u}_2 \mathbf{u}_2^T \mathbf{x} \end{aligned}$$

By setting the latter to a constant c^2 , we can see what shape points $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$ satisfy:

$$\begin{aligned} \lambda_1 \mathbf{x}^T \mathbf{u}_1 \mathbf{u}_1^T \mathbf{x} + \lambda_2 \mathbf{x}^T \mathbf{u}_2 \mathbf{u}_2^T \mathbf{x} &= c \implies \\ \frac{(\mathbf{u}_1 \mathbf{x})^2}{\left(\frac{c}{\sqrt{\lambda_1}}\right)^2} + \frac{(\mathbf{u}_2 \mathbf{x})^2}{\left(\frac{c}{\sqrt{\lambda_2}}\right)^2} &= 1 \end{aligned} \quad (1)$$

Furthermore, 1 says that the ellipse is centred w.r.t. the eigenvectors $\mathbf{u}_1, \mathbf{u}_2$, i.e. that they point to the ellipse axes.

Another way to put this is as follows. Suppose a point on the ellipse is given as a linear combination of unit eigenvectors as basis - \mathbf{u}_1 and \mathbf{u}_2 :

$$\mathbf{x} = y_1 \mathbf{u}_1 + y_2 \mathbf{u}_2 \quad (2)$$

y_1, y_2 are the coordinates of \mathbf{x} w.r.t. the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$, therefore $y_1 = \mathbf{x} \cdot \mathbf{u}_1, y_2 = \mathbf{x} \cdot \mathbf{u}_2$ so Eq. (2) is rewritten as:

$$\frac{y_1^2}{\left(\frac{c}{\sqrt{\lambda_1}}\right)^2} + \frac{y_2^2}{\left(\frac{c}{\sqrt{\lambda_2}}\right)^2} = 1 \quad (3)$$

The points of the latter ellipse are centred w.r.t. the eigenvectors as basis vectors. If Eq. (3) was not centred, there would be a Bxy term in Eq. (3) as we showed in Eq. (2.3). \square

Proof.

Therefore the length of the primary (α) and secondary (β) semi axes is:

$$\alpha = \frac{c}{\sqrt{(\lambda_2)}}, \quad \beta = \frac{c}{\sqrt{(\lambda_1)}}, \quad \lambda_1 \geq \lambda_2 \quad (2.4)$$

Therefore the dominant eigenvalue defines the minor axis (steepest ascent) and the small eigenvalue defines the major axis. \square

EXAMPLE 2.1. Describe the curve (in terms of axes and rotation) $Q(\mathbf{x}) = 8x_1^2 - 4x_1x_2 + 5x_2^2$, $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$.

Solution:

$Q(\mathbf{x})$ can be written as the quadratic form of matrix $\mathbf{A} = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$.

$$Q(\mathbf{x}) = \mathbf{x}^T \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix} \mathbf{x}$$

We want to remove the cross product hence find some new basis axes by making the substitution $\mathbf{x} = \mathbf{P}\mathbf{y}$, where $\mathbf{P} = [\mathbf{u}_1 \quad \mathbf{u}_2]$, where $\mathbf{u}_1, \mathbf{u}_2$ are the eigenvectors of \mathbf{A} .

Verify that

$$\begin{aligned} \lambda_1 = 9 : \mathbf{u}_1 &= \begin{bmatrix} -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}^T, \quad \lambda_2 = 4 : \mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}^T \\ \therefore \mathbf{P} &= \frac{1}{\sqrt{5}} \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix} \end{aligned}$$

We've shown that \mathbf{P} diagonalises \mathbf{A} , i.e. $\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{D}$ therefore using $\mathbf{x} = \mathbf{P}\mathbf{y}$ we end up with

$$\begin{aligned} \mathbf{y}^T \mathbf{D} \mathbf{y} &= 1 \implies \\ 9y_1^2 + 4y_2^2 &= 1 \implies \\ \frac{y_1^2}{\left(\frac{1}{3}\right)^2} + \frac{y_2^2}{\left(\frac{1}{2}\right)^2} &= 1 \end{aligned} \quad (\clubsuit)$$

From Eq. (2.4), the length of the major and minor axes are respectively $\alpha = 1/2, \beta = 1/3$. Note that we have just verified that $\alpha = 1/\sqrt{\lambda_2}, \beta = 1/\sqrt{\lambda_1}, \lambda_1 \geq \lambda_2$.

We finally want to know how the axes y_1, y_2 of ellipse Eq. (♣), so essentially the eigenvectors of matrix \mathbf{A} are rotated w.r.t. the standard axes. The angle of u_1 is $\arctan(-0.5) = -26.56 \text{ mod } 180^\circ = 153.44^\circ$ and the angle of u_2 is $\arctan(2) = 63.43^\circ \text{ (mod } 180 \text{ because the angle is relative to the positive } x \text{ axis)}$. The ellipse with its eigenvectors is shown below.

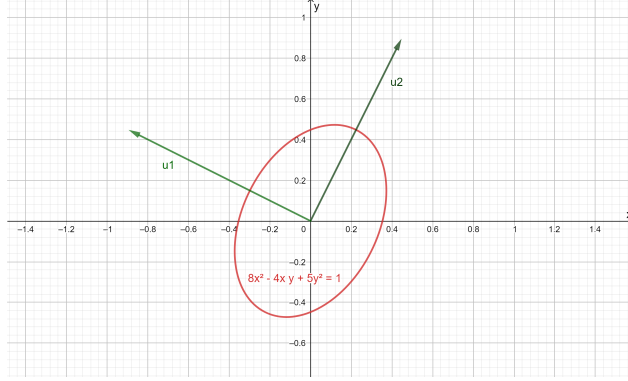


Figure 6: The ellipse given in the example. Notice how the unit eigenvectors point to the axes.

3 Classifying quadratic surfaces

As already mentioned, when \mathbf{A} is a $n \times n$ matrix, the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ is a real-valued function $Q : \mathbb{R}^n \rightarrow \mathbb{R}$. For example, if the domain is \mathbb{R}^2 , then for points (x_1, x_2, z) for a certain value $z = Q(x_1, x_2)$ correspond a horizontal cut of $z = Q(x_1, x_2)$, therefore either to an ellipse or hyperbola or their degenerate cases [1]. As shown in the figures below, $z = Q(x_1, x_2)$ describes

- an elliptic paraboloid,
- or an elliptic hyperboloid

For the equations of those shapes, see Appendix 2. The important property is that an elliptic paraboloid is either convex or concave, therefore if its vertex is at $(0, 0)$, it will be either non-negative or non-positive. Some examples are illustrated in Fig. 8.

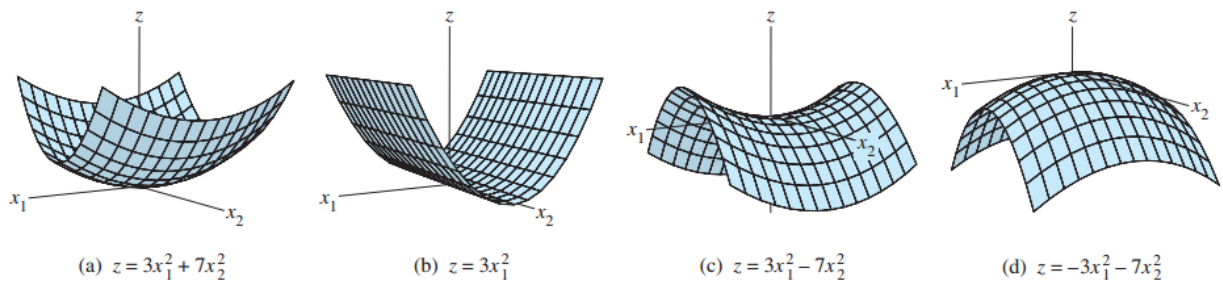


Figure 7: The ellipse given in the example. Notice how the unit eigenvectors point to the axes [1].

DEFINITION 3.1. A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ is

- positive definite (PD) if $Q(\mathbf{x}) > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$.
- negative definite (ND) if $Q(\mathbf{x}) < 0 \quad \forall \mathbf{x} \neq \mathbf{0}$.
- semi-positive (semi-negative) definite (SPD, SND) if $Q(\mathbf{x}) \geq 0 \text{ (} \leq 0 \text{)} \quad \forall \mathbf{x} \neq \mathbf{0}$.

■ *indefinite if $Q(\mathbf{x})$ assumes both positive and negative values.*

☞ **Comment:** Often, instead of saying that the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is positive definite, we instead say that \mathbf{A} is positive definite. Both have the same meaning.

EXAMPLE 3.1. *Prove that if \mathbf{A} and \mathbf{B} are positive definite, then so is $\mathbf{A} + \mathbf{B}$.*

Solution:

$$\begin{aligned}\mathbf{x}^T \mathbf{A} \mathbf{x} &> 0 \quad \forall \mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\} \\ \mathbf{x}^T \mathbf{B} \mathbf{x} &> 0 \quad \forall \mathbf{x} \in \mathbb{R}^n - \{\mathbf{0}\} \\ \therefore \mathbf{x}^T (\mathbf{A} + \mathbf{B}) \mathbf{x} &= \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{B} \mathbf{x} > 0\end{aligned}$$

■

Looking at Fig. 7, forms (a), (b) are semi-positive definite but (a) is described better at positive negative, as it is only 0 at the origin. Positive definite and symmetric matrices are useful in a number of fields, such as statistics. We will prove how the following statements are equivalent [3].

- (a) \mathbf{A} is positive definite and symmetric.
- (b) All eigenvalues are positive.
- (c) There exists factorisation of $\mathbf{A} = \mathbf{R}^T \mathbf{R}$.
- (d) All upper left determinants, including the main determinant itself, are positive.
- (e) All pivots (without row exchanges) satisfy $d_k > 0$.

By upper left detetminants, we mean that if e.g. $\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ then they are 2, $\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3$, $\begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 4$.

LEMMA 3.1. (a) \Leftrightarrow (b) *Let \mathbf{A} be an $n \times n$ symmetric matrix. Then its quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is:*

- *positive definite iff the eigenvalues of \mathbf{A} are all positive.*
- *negative definite iff the eigenvalues of \mathbf{A} are all negative.*
- *indefinite iff \mathbf{A} has both positive and negative eigenvalues.*

Proof. (b \Rightarrow a)

Assume all eigenvalues λ_i , $1 \leq i \leq n$ are positive. \mathbf{A} is symmetric therefore by the Principal Axes theorem, there exists a change of variable $\mathbf{x} = \mathbf{P} \mathbf{y}$ such that:

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{D} \mathbf{y} = \lambda_1 \mathbf{y}_1^2 + \lambda_2 \mathbf{y}_2^2 + \dots + \lambda_n \mathbf{y}_n^2$$

So clearly $Q(\mathbf{x}) > 0$.

(a \Rightarrow b)

Assume that \mathbf{A} is positive definite, i.e. $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \quad \forall \quad \mathbf{x} \neq \mathbf{0}$. If λ is an eigenvalue and \mathbf{u} its eigenvector, then:

$$\begin{aligned}\mathbf{A} \mathbf{u} &= \lambda \mathbf{u} \implies \\ \mathbf{u}^T \mathbf{A} \mathbf{u} &= \lambda \mathbf{u}^T \mathbf{u} \implies \\ \lambda &= \frac{\mathbf{u}^T \mathbf{A} \mathbf{u}}{\mathbf{u}^T \mathbf{u}}\end{aligned}$$

$\mathbf{u}^T \mathbf{A} \mathbf{u} > 0$ because of the hypothesis (zero eigenvectors are non meaningful) therefore $\lambda > 0$. □

Positive definite matrices also admit a factorisation called Cholesky factorisation. How we obtain is explained in Appendix A.3.

LEMMA 3.2. (a \Leftrightarrow c) *A symmetric and positive definite matrix \mathbf{A} admits a factorisation $\mathbf{A} = \mathbf{R}^T \mathbf{R}$, where the columns of \mathbf{R} are linearly independent and vice versa.*

Proof.

Appendix A.3. □

LEMMA 3.3. (a) \Rightarrow (d) If \mathbf{A} is positive definite, then all upper left determinants are positive.

Proof.

Before we start, recall the following about the determinant of any matrix.

■ Fact 1. The determinant of any matrix is the product of its eigenvalues.

Let \mathbf{A} be positive definite, $\mathbf{x} \in \mathbb{R}^n$, $k \leq n$, and \mathbf{x}_k the first k entries of \mathbf{x} . Then $\mathbf{x} = \begin{bmatrix} \mathbf{x}_k & \mathbf{0} \end{bmatrix}$. \mathbf{A} is symmetric therefore partitioning its first $k \times k$ elements with the rest, we can write it as $\mathbf{A} = \begin{bmatrix} \mathbf{A}_k & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix}$. We know that

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &> 0 \implies \\ \begin{bmatrix} \mathbf{x}_k^T & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{A}_k & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k & \mathbf{0} \end{bmatrix} &> 0 \implies \\ \mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_k &> 0 \quad \forall \quad \mathbf{x}_k \neq \mathbf{0} \end{aligned}$$

Therefore ((a) \Rightarrow (b)) for all eigenvalues of \mathbf{A} :

$$\begin{aligned} \lambda_1, \lambda_2, \dots, \lambda_n &> 0 \implies \\ \lambda_1, \lambda_2, \dots, \lambda_k &> 0 \implies \\ \det(\mathbf{A}_k) &= \lambda_1 \lambda_2 \dots \lambda_k > 0 \quad \forall \quad k \leq n \end{aligned}$$

□

LEMMA 3.4. (a) \Rightarrow (e) All pivots d_k of a real symmetric positive definite matrix satisfy $d_k > 0$.

Proof. As shown in Appendix A.3, we can decompose \mathbf{A} as a product of a (lower triangular) \times (diagonal) \times (upper triangular) matrices, i.e. "LDU":

$$\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{U}$$

\mathbf{L}, \mathbf{U} are unit triangular and \mathbf{D} is diagonal containing the pivots. As in the previous lemma, partition \mathbf{A} as $\mathbf{A} = \begin{bmatrix} \mathbf{A}_k & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix}$ and $\mathbf{L}, \mathbf{D}, \mathbf{U}$ in the first $k \times k$ elements:

$$\begin{aligned} \begin{bmatrix} \mathbf{A}_k & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} &= \begin{bmatrix} \mathbf{L}_k & \mathbf{0} \\ * & * \end{bmatrix} \begin{bmatrix} \mathbf{D}_k & \mathbf{0} \\ \mathbf{0} & * \end{bmatrix} \begin{bmatrix} \mathbf{U}_k & * \\ * & \mathbf{0} \end{bmatrix} \implies \\ \mathbf{A}_k &= \mathbf{L}_k \mathbf{D}_k \mathbf{U}_k \end{aligned}$$

Taking determinants above and because \mathbf{L}, \mathbf{U} are triangular and have ones along the diagonal, while \mathbf{D} has the pivots:

$$\begin{aligned} \det(\mathbf{A}_k) &= \det(\mathbf{D}_k) = d_1 d_2 \dots d_k \\ \therefore d_k &= \frac{\det(\mathbf{A}_k)}{\det(\mathbf{A}_{k-1})} > 0, \quad k = 1, 2, \dots, n \end{aligned}$$

□

LEMMA 3.5. (e) \Rightarrow (a) If a real symmetric matrix \mathbf{A} has only positive pivots, then it is positive definite.

Proof. We have seen in App. A.3 that when \mathbf{A} is symmetric, the LDU decomposition is:

$$\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{L}^T$$

, where \mathbf{D} stores the pivots. If $d_{ii} > 0$, then

$$\mathbf{L} \mathbf{D} \mathbf{L}^T = \mathbf{L} \mathbf{D}^{1/2} \mathbf{D}^{1/2} \mathbf{L}^T = \mathbf{R}^T \mathbf{R}, \quad \mathbf{R} = \mathbf{D}^{1/2} \mathbf{L}^T$$

Therefore

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{R}^T \mathbf{R} \mathbf{x} = (\mathbf{R} \mathbf{x})^T (\mathbf{R} \mathbf{x}) = \|\mathbf{R} \mathbf{x}\|^2 > 0$$

□

We have just proven the following important theorem. This is called Sylvester's theorem and various other interesting proofs are found in [4] and [5].

THEOREM 3.1. *Each of the following tests is a necessary and sufficient condition for the real symmetric matrix \mathbf{A} to be positive definite [3].*

1. $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \quad \forall \mathbf{x} \neq \mathbf{0}$
2. All the eigenvalues of A satisfy $\lambda_i > 0$.
3. \mathbf{A} admits a decomposition $\mathbf{R}^T \mathbf{R}$.
4. All the upper left submatrices \mathbf{A}_k have positive determinants
5. All the pivots (without row exchanges) satisfy $d_k > 0$

EXAMPLE 3.2. *Prove that every positive definite matrix is invertible.*

Solution:

We know

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \dots \lambda_n$$

If \mathbf{A} wasn't invertible then:

$$\exists \lambda_i : \det(\mathbf{A}) = 0 \Leftrightarrow \lambda_i = 0, \quad 1 \leq i \leq n$$

That's contradicts the fact that because \mathbf{A} positive definite, then $\lambda_i > 0$.

EXAMPLE 3.3. *Is $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ positive definite [1]?*

Solution:

It may look like it is positive definite, however we need to find \mathbf{A} 's eigenvalues and write it as $\mathbf{y}^T \mathbf{D} \mathbf{y}$ first. The matrix \mathbf{A} is:

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

We find $\lambda_1 = 5, \lambda_2 = 2, \lambda_3 = -1$. Therefore Q is indefinite. ■

EXAMPLE 3.4. *Determine whether matrix $\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ is positive definite [3].*

Solution:

First way:

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} &= 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 \\ &= 2\left(x_1 - \frac{1}{2}x_2\right)^2 + \frac{3}{2}\left(x_2 - \frac{2}{3}x_3\right)^2 + \frac{4}{3}x_3^2 > 0 \end{aligned}$$

Second way:

$$\lambda = 2, 2 \pm \sqrt{2}$$

Third way:

$$A_1 = 2$$

$$A_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3$$

$$A_3 = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 4$$

$$\therefore A_1, A_2, A_3 > 0$$

Fourth way:

The row reduced Echelon form of \mathbf{A} is \mathbf{I}_3 (3×3 unit matrix) so the pivots are 1, 1, 1 - all positive.

Finally, we have some similar criteria for a positive semi-definite matrix. Inverting the signs wouldn't work if we were to guess the criteria for a negative definite matrix but this will not be covered in this article.

EXAMPLE 3.5. Let \mathbf{A} be an $n \times n$ invertible matrix. Show that if the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is positive, then so is $\mathbf{x}^T \mathbf{A}^{-1} \mathbf{x}$ and that the inverse is symmetric..

Solution 1:

We know from Theorem 3.1 that if $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ $\mathbf{x} \neq \mathbf{0}$, then all its eigenvalues λ_i are positive. For the eigenvalues of the inverse we know:

$$\lambda_{-1,i} = \frac{1}{\lambda_i} > 0$$

From the same theorem, we have just proven what is required. Also:

$$(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} = \mathbf{A}^{-1}$$

Solution 2:

Define $\mathbf{y} = \mathbf{A} \mathbf{x}$ and take the quadratic form of \mathbf{A}^{-1} :

$$\begin{aligned} \mathbf{y}^T \mathbf{A}^{-1} \mathbf{y} &= \mathbf{x}^T \mathbf{A}^T \mathbf{A}^{-1} \mathbf{A} \mathbf{x} \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{x} \\ &= \mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \end{aligned}$$

THEOREM 3.2. Each of the following tests is a necessary and sufficient condition for the real symmetric matrix \mathbf{A} to be positive semi-definite.

1. $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n$
2. All the eigenvalues of A satisfy $\lambda_i \geq 0$.
3. A admits a decomposition $\mathbf{R}^T \mathbf{R}$.
4. All the upper left submatrices \mathbf{A}_k have non-negative determinants
5. All the pivots (without row exchanges) satisfy $d_k \geq 0$

EXAMPLE 3.6. [6] Find the values of b such that the following matrix is semi-positive definite.

$$\begin{bmatrix} 2 & -1 & b \\ -1 & 2 & -1 \\ b & -1 & 2 \end{bmatrix}$$

Solution:

$$\begin{vmatrix} 2 & -1 & b \\ -1 & 2 & -1 \\ b & -1 & 2 \end{vmatrix} = 6 - 2 + b + b(1 - 2b) \geq 0 \Leftrightarrow -1 \leq b \leq 2,$$

$$\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} > 0,$$

$$2 > 0$$

Therefore $-1 \leq b \leq 2$. ■

For reference, we also have the following criterion for negative definiteness. The proof is omitted, however various proofs are analysed at [5].

THEOREM 3.3. Each of the following tests is a necessary and sufficient condition for the real symmetric matrix \mathbf{A} to be ND.

1. $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n$
2. All of the eigenvalues of \mathbf{A} satisfy $\lambda_i < 0$
3. All of its k -th order upper left principal minor shave sign $(-1)^k$.
4. All the pivots (without row exchanges) satisfy $d_k < 0$

4 Constrained optimisation

4.1 Optimisation of quadratic form in the unit sphere

Often we want to minimise or maximise a quadratic form subject to a constrained. One such constraint that allows us to easily study that is that the coordinate vector is unit, i.e. $\|\mathbf{x}\| = 1$, or equivalently

$$\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x} = x_1^2 + x_2^2 + \dots + x_n^2 = 1 \quad (4.1)$$

We will also see that when a quadratic form has no cross-product terms, it is easier to find its extrema for $\|\mathbf{x}\| = 1$.

EXAMPLE 4.1. Find the extrema of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subject to $\|\mathbf{x}\| = 1$ [1].

Solution:

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\leq 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

Notice that $Q(\mathbf{x}) = 9$ for $\mathbf{x} = (1, 0, 0)$. Therefore the maximum value is 9. Similarly,

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\geq 3x_1^2 + 3x_2^2 + 3x_3^2 \\ &= 3(x_1^2 + x_2^2 + x_3^2) \\ &= 3 \end{aligned}$$

, with $Q(\mathbf{x}) = 3$ for $\mathbf{x} = (0, 0, 1)$, therefore the minimum is 3. ■

Let's consider matrix $\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$ and its quadratic form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$, $\mathbf{x} \in \mathbb{R}^2$. We want to visualise the extrema of $Q(\mathbf{x})$ given the constraint $\|\mathbf{x}\| = 1$.

By setting $z = \mathbf{x}^T \mathbf{A} \mathbf{x} = 3x_1^2 + 7x_2^2$, it is clear that the surface $z(x_1, x_2)$ is a positive paraboloid with minimum of 0 at $(0, 0)$. The condition $\|\mathbf{x}\| = 1$ adds the constraint that the points for which $z(x_1, x_2)$ is minimised or maximised must lie on the cylinder along the z axis with radius of 1. Note also that $z(x_1, x_2)$ grows fastest in height along axis x_2 and slowest along x_1 .

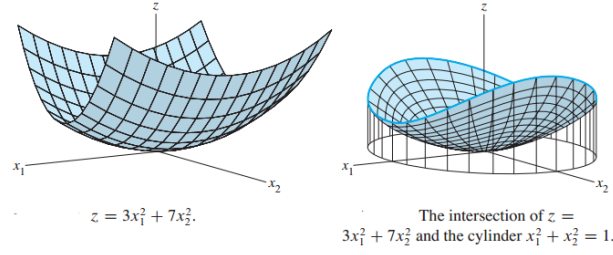


Figure 8: The surface and its intersection with the cylinder $\|\mathbf{x}\| = 1$, representing the constraint [1].

The highest value of 7 given $\|\mathbf{x}\| = 1$ occurs at the eigenvector $\mathbf{u}_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and it's equal to the eigenvalue $\lambda_1 = 7$. The minimum value of 3 occurs at the eigenvector $\mathbf{u}_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and it's equal to the eigenvalue $\lambda_2 = 3$.

LEMMA 4.1. *If \mathbf{A} is a symmetric matrix, all possible values for $\mathbf{x}^T \mathbf{A} \mathbf{x}$, given $\|\mathbf{x}\| = 1$, are real.*

Proof. **TODO: both in one proof** □

LEMMA 4.2. *Denote the right and left endpoints of this interval by m and M respectively, i.e.*

$$m = \min\{\mathbf{x}^T \mathbf{A} \mathbf{x}, \|\mathbf{x}\| = 1\}, \quad M = \max\{\mathbf{x}^T \mathbf{A} \mathbf{x}, \|\mathbf{x}\| = 1\} \quad (4.2)$$

If λ is an eigenvalue of \mathbf{A} , then $m \leq \lambda \leq M$.

Proof.

Let \mathbf{u} be a unit eigenvector of \mathbf{A} for the eigenvalue λ and m, M the minimum and maximum of the quadratic form, then:

$$\begin{aligned} \mathbf{u}^T \mathbf{A} \mathbf{x} &= \mathbf{u}^T \lambda \mathbf{u} = \lambda \|\mathbf{u}\|^2 = \lambda \\ \therefore m &\leq \lambda \leq M \end{aligned}$$
□

The next theorem says that there exists eigenvalues of \mathbf{A} m and M are themselves.

THEOREM 4.1. *Let \mathbf{A} be a symmetric matrix and m, M be defined as in Eq. (4.2). Then M is equal to the greatest eigenvalue λ_1 of \mathbf{A} and m to the smallest eigenvalue λ_n . $\mathbf{x}^T \mathbf{A} \mathbf{x} = M$ when $\mathbf{x} = \mathbf{u}_1$ and $\mathbf{x}^T \mathbf{A} \mathbf{x} = m$ when $\mathbf{x} = \mathbf{u}_n$, where $\mathbf{u}_1, \mathbf{u}_n$ are the unit eigenvectors corresponding to λ_1, λ_n [1].*

Proof.

Before we start, we use the following facts:

1. If \mathbf{A} is orthogonally diagonalised as $\mathbf{P} \mathbf{D} \mathbf{P}^{-1}$, then by setting $\mathbf{x} = \mathbf{P} \mathbf{y}$ the cross-product is removed from the quadratic form:

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{y}^T \mathbf{A} \mathbf{y} \quad (\spadesuit)$$

2. $\mathbf{x}^T \mathbf{A} \mathbf{x}$ and $\mathbf{y}^T \mathbf{D} \mathbf{y}$ assume the same range of values. The transformation $\mathbf{x} = \mathbf{P} \mathbf{y}$ only rotates \mathbf{x} without changing its magnitude, as \mathbf{P} is orthonormal, i.e. $\mathbf{P}^T \mathbf{P} = \mathbf{I}$.

$$\begin{aligned} \|\mathbf{x}\|^2 &= \|\mathbf{P} \mathbf{y}\|^2 = 1 \implies \\ (\mathbf{P} \mathbf{y})^T (\mathbf{P} \mathbf{y}) &= \mathbf{y}^T \mathbf{P}^T \mathbf{P} \mathbf{y} = 1 \implies \\ \|\mathbf{y}\| &= 1 \end{aligned} \quad (\spadesuit\spadesuit)$$

3. The eigenvalues of a real and symmetric matrix are non-negative.

To simplify the notation, assume \mathbf{A} is a 3×3 matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and arrange the corresponding eigenvectors in \mathbf{P} in the same order so that:

$$\mathbf{P} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3], \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix}$$

For any vector $\mathbf{y} \in \mathbb{R}^3$:

$$\begin{aligned} \lambda_1 y_1^2 &= \lambda_1 y_1^2 \\ \lambda_2 y_2^2 &= \lambda_1 y_2^2 \\ \lambda_3 y_3^2 &= \lambda_1 y_3^2 \implies \\ \mathbf{y}^T \mathbf{D} \mathbf{y} &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 \\ &\leq \lambda_1 y_1^2 + \lambda_1 y_2^2 + \lambda_1 y_3^2 \\ &= \lambda_1 (y_1^2 + y_2^2 + y_3^2) \\ &= \lambda_1 \|\mathbf{y}\|^2 = \lambda_1 \end{aligned}$$

Therefore $\lambda_1 \geq M$. Also $\mathbf{e}_1^T \mathbf{D} \mathbf{e}_1 = \lambda_1$ so the maximum of $\mathbf{y}^T \mathbf{A} \mathbf{y}$ is λ_1 attained for $\mathbf{y} = \mathbf{e}_1$. Similarly, λ_3 is the minimum of $\mathbf{y}^T \mathbf{A} \mathbf{y}$ for $\mathbf{y} = \mathbf{e}_3$.

From Fact 2 (**TODO: re-do the numbering above**) we know the two quadratic forms assume the same range, therefore when one is maximum the other is maximum too. Hence $M = \lambda_1 = \mathbf{e}_1^T \mathbf{D} \mathbf{e}_1 = \mathbf{u}_1^T \mathbf{A} \mathbf{u}_1$, so the maximum is attained when $\mathbf{x} = \mathbf{P} \mathbf{e}_1 = \mathbf{u}_1$. $m = \lambda_3 = \mathbf{e}_3^T \mathbf{D} \mathbf{e}_3 = \mathbf{u}_3^T \mathbf{A} \mathbf{u}_3$ attained when $\mathbf{x} = \mathbf{P} \mathbf{e}_3 = \mathbf{u}_3$. \square

EXAMPLE 4.2. Let $\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Find the maximum value of the quadratic form $\mathbf{x}^T \mathbf{A} \mathbf{x}$ subject to $\|\mathbf{x}\| = 1$ [1].

Solution:

We find the greatest eigenvalue to be $\lambda = 6$.

The constrained maximum of $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is attained when \mathbf{x} is the unit eigenvector corresponding to $\lambda = 6$. Solving $\mathbf{A} - 6\mathbf{I} = 0$ we find $\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$ therefore $\mathbf{u} = 1/\sqrt{3} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$. \blacksquare

Below we maximise the quadratic form with some additional constraints.

THEOREM 4.2. [1] Let $\lambda_1, \lambda_2, \dots, \lambda_n$, $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ be the eigenvalues and eigenvectors of a symmetric matrix \mathbf{A} . Then the maximum value of $\mathbf{x}^T \mathbf{A} \mathbf{x}$ subject to the constraints

$$\mathbf{x}^T \mathbf{u}_1 = 0, \quad \|\mathbf{x}\| = 1$$

is the second greatest eigenvalue λ_2 and it is attained when \mathbf{x} is the unit eigenvector \mathbf{u}_2 corresponding to λ_2 .

As a generalisation, we have the following.

THEOREM 4.3. [1] Let \mathbf{A} be a symmetric $n \times n$ matrix with an orthogonal diagonalisation $\mathbf{A} = \mathbf{P} \mathbf{D} \mathbf{P}^{-1}$, where the entries on the diagonal are arranged such that $\lambda_1 \geq \lambda_2 \dots, \lambda_n$ and the columns of \mathbf{P} are the corresponding eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$. Then for $k = 1, 2, \dots, n$ the maximum value of $\mathbf{x}^T \mathbf{A} \mathbf{x}$ subject to

$$\mathbf{x}^T \mathbf{u}_1 = 0, \mathbf{x}^T \mathbf{u}_2 = 0, \dots, \mathbf{x}^T \mathbf{u}_n = 0, \quad \|\mathbf{x}\| = 1$$

is the eigenvalue λ_k , and this maximum is attained for $\mathbf{x} = \mathbf{u}_k$.

EXAMPLE 4.3. Maximise $q(x, y) = xy$ given the constraint $4x^2 + 9y^2 \leq 36$ [1].

Note: Each point (x, y) that satisfies the constraint is called feasible solution. The curve $4x^2 + 9y^2 = 36$ is called constraint curve. The set of points for which $4x^2 + 9y^2 = \text{const}$ is called indifference curve.

Solution:

In our problem, we do not have a "unit sphere constraint" for the input free variables so we need to create one. We have

$$\begin{aligned} 4x^2 + 9y^2 &\leq 36 \implies \\ \left(\frac{x}{3}\right)^2 + \left(\frac{y}{2}\right)^2 &\leq 1 \xrightarrow[y'=\frac{y}{2}]{x'=\frac{x}{3}} \\ (x')^2 + (y')^2 &\leq 1 \end{aligned}$$

The function we want to minimise becomes

$$q(x', y') = 6x'y'$$

It can now be written in quadratic form given the "unit sphere constraint" therefore its maximum can be found.

$$q(x', y') = 6x'y' = \mathbf{x} \begin{bmatrix} 0 & 3 \\ 0 & 3 \end{bmatrix} \mathbf{x}$$

For matrix \mathbf{A} , $\lambda = 3, -3$ with eigenvectors $\mathbf{x} = [1/\sqrt{2} \quad 1/\sqrt{2}]^T$, $[-1/\sqrt{2} \quad 1/\sqrt{2}]^T$ Therefore the maximum value of 3 is attained at $(x', y') = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ therefore for $(x, y) = (\frac{3}{\sqrt{2}}, \frac{2}{\sqrt{2}})$.



4.2 (Optional) Optimisation of general 2nd order form

4.2.1 Unbounded quadric problem (part 1)

In this section, we want to minimise an objective function of the form

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0 \quad (4.3)$$

In order to minimise the function in Eq. (4.3), we will show that \mathbf{H} has to be semi positive definite. Assume that \mathbf{H} is negative definite, i.e. there exists $\mathbf{d} \in \mathbb{R}^n$ such that $\mathbf{d}^T \mathbf{H} \mathbf{d} < 0$, then:

$$f(t\mathbf{x}) = \frac{1}{2} t^2 \mathbf{d}^T \mathbf{H} \mathbf{d} + t \mathbf{c}^T \mathbf{d} + c_0 \rightarrow -\infty \quad \text{as } t \rightarrow \infty$$

Therefore \mathbf{H} has to be positive definite and according to Appendix A.3, it admits an "LDU" factorisation, i.e. $\mathbf{H} = \mathbf{L}^T \mathbf{D} \mathbf{L}$ with \mathbf{L} lower triangular, where $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$, with $d_i \geq 0 \quad \forall \quad 1 \leq i \leq n$ according to Theorem 3.2. As shown in [7], The problem becomes

$$\begin{aligned} &\text{minimise } \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0 \\ &= \text{minimise } \frac{1}{2} \mathbf{x}^T \mathbf{L}^T \mathbf{D} \mathbf{L} \mathbf{x} + \mathbf{c}^T (\mathbf{L}^T)^{-1} \mathbf{L}^T \mathbf{x} + c_0 \\ &= \text{minimise } \sum_{k=1}^n \frac{1}{2} d_k y_k^2 + \tilde{c}_k y_k + c_0 \\ &= c_0 + \sum_{k=1}^n \text{minimise } \frac{1}{2} d_k y_k^2 + \tilde{c}_k y_k \end{aligned}$$

We have introduced the new coordinates

$$\mathbf{y} = \mathbf{L}^T \mathbf{x}, \quad \tilde{\mathbf{c}} = \mathbf{L}^{-1} \mathbf{c} \quad (4.4)$$

The minimisation can now be decoupled to n independent scalar quadratic problems.

4.2.2 Scalar quadratic optimization without constraints

The scalar quadratic form

$$p(x) = ax^2 + 2bx + c$$

(2 factor because it's generated by a symmetric matrix) has a minimum only if $a > 0$. By differentiating we have:

$$p'(x) = 2ax + 2b$$

$$p'(x) \geq 0 \Leftrightarrow x \geq -\frac{b}{a}$$

$$p'(x) = 0 \Leftrightarrow x = -\frac{b}{a}$$

Therefore $p(x)$ is strictly decreasing for $x < -\frac{b}{a}$ and strictly increasing for $x > -\frac{b}{a}$, which means that the minimum occurs at

$$x^* = -\frac{b}{a}, \text{ where } p(x^*) = c - \frac{b^2}{a} \quad (4.5)$$

4.2.3 Unbounded quadratic problem (part 2)

Applying Eq. (4.5) to solve our minimisation sub-problems separately, we find for each \hat{y}_i that $\hat{\mathbf{y}}_i = -\frac{\tilde{c}_k}{d_k}$, given the semi-positive definiteness condition $d_k \geq 0$. So for a minimiser $\hat{\mathbf{y}} = [\hat{y}_1 \ \hat{y}_2 \ \dots \ \hat{y}_n]^T$ we have the following lemma [7].

LEMMA 4.3. *The objective function*

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0 \quad (4.6)$$

is minimised when

$$g(\mathbf{x}) = \sum_{k=1}^n \frac{1}{2} d_k y_k^2 + \tilde{c}_k y_k$$

is minimised, where $\mathbf{x} = \mathbf{L} \mathbf{D} \mathbf{L}^T$, $\mathbf{y} = \mathbf{L}^T \mathbf{x}$, $\tilde{\mathbf{c}} = \mathbf{L}^{-1} \mathbf{c}$. Vector $\hat{\mathbf{y}}$ is a minimiser of $g(x)$ if:

1. $d_i \geq 0 \ \forall \ 1 \leq i \leq n \Leftrightarrow \mathbf{H}$ is semi positive definite.
2. $d_k \hat{y}_k = -\tilde{c}_k \Leftrightarrow \mathbf{D} \mathbf{L}^T \hat{\mathbf{x}} = -\mathbf{L}^{-1} \mathbf{c} \Leftrightarrow \mathbf{L} \mathbf{D} \mathbf{L}^T \hat{\mathbf{x}} = -\mathbf{c}$

\hat{x} is simply the respective $\hat{\mathbf{x}}$ vector for which \mathbf{y} is a (unique) minimiser. Using LDU factorisation and for the global minimiser \mathbf{x}^* , constraint (2) can be simplified to:

$$\mathbf{H} \mathbf{x}^* = -\mathbf{c} \xLeftrightarrow{\mathbf{H} \text{ PD}} \mathbf{x}^* = -\mathbf{H}^{-1} \mathbf{c} \quad (4.7)$$

The second form holds only if \mathbf{H} is positive definite. Then the minimum value of $f(\mathbf{x})$ is

$$f(-\mathbf{H}^{-1} \mathbf{c}) = -\frac{1}{2} \mathbf{c}^T \mathbf{H}^{-1} \mathbf{c} + c_0$$

Note that to remember Eq. (4.7), we can compare the objective function $f(\mathbf{x})$ in Eq. (4.6) with the scalar quadratic form $f(x) = ax^2 + 2bx + c$ and the minimiser in Eq. (4.7) with the scalar one $ax = -b$.

Furthermore, inverting any of the conditions in Lemma 4.3, we can find when $f(x) = \frac{1}{2}\mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0$ is unbounded from below [7].

LEMMA 4.4. $f(x) = \frac{1}{2}\mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0$ is unbounded from below if:

1. Some $d_k < 0 \Leftrightarrow \mathbf{H}$ is not semi-positive definite.
2. There exists k such that one of the equations

$$d_k \hat{y}_k = -\tilde{c}_k$$

do not have a solution, i.e. $d_k = 0$ and $\tilde{c}_k \neq 0$ $d_k \hat{y}_k = -\tilde{c}_k \Leftrightarrow \mathbf{D} \mathbf{L}^T \hat{\mathbf{x}} = -\mathbf{L}^{-1} \mathbf{c}$ i.e. the last condition says that the system $\mathbf{H} \hat{\mathbf{x}} = -\mathbf{c}$ does not have a solution.

The above arguments can be formalised as the following theorem

THEOREM 4.4. Let $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{c}^T\mathbf{x} + c_0$. Then \mathbf{x}^* is a minimiser if

1. \mathbf{H} is semi positive definite.
2. $\mathbf{H}\mathbf{x}^* = -\mathbf{c}$

Note: The condition that $\mathbf{H}\mathbf{x}^* = -\mathbf{c}$ has a solution equivalently tells us that $\mathbf{c} \in \mathcal{R}(\mathbf{H})$, where \mathcal{R} denotes the range, also known as column space.

EXAMPLE 4.4. Minimise the quadratic form $Q(x_1, x_2) = 4x_1^2 - 2x_1x_2 + 3x_2^2 + 3x_1 - 2x_2 + 1$, $x_1, x_2 \in \mathbb{R}$ [8].

Solution:

Write $Q(x_1, x_2)$ in the general quadratic form as in Eq. (4.3).

$$Q(x_1, x_2) = \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 8 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} + 1$$

Therefore $\mathbf{H} = \begin{bmatrix} 8 & -2 \\ -2 & 6 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$ (don't forget to factor out the $\frac{1}{2}$). Then we have showed that the minimiser vector \mathbf{x}^* is found by solving the system

$$\begin{bmatrix} 8 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

\mathbf{H} is positive definite thus invertible therefore $f(\mathbf{x})$ really has a minimum, obtained by Gaussian elimination at

$$\mathbf{x}^* = \begin{bmatrix} -\frac{7}{22} \\ \frac{5}{22} \end{bmatrix} \approx \begin{bmatrix} -.318 \\ .227 \end{bmatrix}$$

with value $Q(\mathbf{x}^*) = Q(-\frac{7}{22}, \frac{5}{22}) \approx .295$ ■.

4.2.4 Unconstrained optimisation using analytic methods; the Hessian matrix

We will show how we can minimise the function in the last example using analytical methods. Algebraic and analytic methods are connected by the Hessian matrix \mathbf{H} , which appears in both. We can often use analytical methods to verify our results. First a reminder.

DEFINITION 4.1. The gradient or "del" of a multi variable differentiable function $f(x_1, x_2, \dots, x_n)$ is a vector operator denoted as ∇f or simply ∇f . In 3D cartesian coordinates, the gradient is given by the vector

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} \quad (4.8)$$

Alternatively, in more compact notation and denoting $\partial_x f := \frac{\partial f}{\partial x}$, it is given by $\nabla f = (\partial_x f, \partial_y f, \partial_z f)$ or in matrix notation $\nabla f = [\partial_x f \quad \partial_y f \quad \partial_z f]^T$. Some properties of the gradient are presented in A.4.

The gradient is useful because it lets us compute the Taylor series of a multi variable function, which in terms lets us derive an optimisation property of the Hessian matrix.

The following definitions help us establish a criterion for multi variable function optimisation.

DEFINITION 4.2. A differentiable multi variable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ exhibits a critical point at $\mathbf{x}^* = (x_1, x_2, \dots, x_n)$ if:

$$\nabla f(\mathbf{x}^*) = 0 \Leftrightarrow f_{x_1}(\mathbf{x}^*) = f_{x_2}(\mathbf{x}^*) = \dots = f_{x_n}(\mathbf{x}^*) = 0 \quad (4.9)$$

DEFINITION 4.3. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function whose second partial derivatives exist and are continuous over the domain of the function. Its Hessian matrix \mathbf{H} is a $n \times n$ matrix defined as

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \quad (4.10)$$

Hessian in Eq. (4.10) can tell us whether the critical point is a local maximum, minimum, or saddle point. For a 2-variable function, we have proven the following theorem in App. A.5 using Taylor series. A more detailed proof is found in [9].

THEOREM 4.5. *Let $f(x, y)$ be a twice differentiable function, and let (x_0, y_0) be a critical point for f , i.e. $\nabla f(x_0, y_0) = 0$.*

1. *If $\mathbf{H}_f(x_0, y_0)$ is positive definite, then x_0, y_0 is a local minimum for f .*
2. *If $\mathbf{H}_f(x_0, y_0)$ is negative definite, then x_0, y_0 is a local maximum for f .*
3. *If $\mathbf{H}_f(x_0, y_0)$ is indefinite, then x_0, y_0 is a saddle point for f .*
4. *If $\mathbf{H}_f(x_0, y_0) = 0$, then the second derivative test doesn't tell us anything and we need a higher order approximation.*

We can use Theorem 4.5 to solve the example in the previous section using analytic ways.

EXAMPLE 4.5. *Minimise the quadratic form $Q(x, y) = 4x^2 - 2xy + 3y^2 + 3x - 2y + 1$, $x, y \in \mathbb{R}$.*

Solution:

Compute the partial derivatives, denoting $f_x := \frac{\partial f}{\partial x}$, $f_y := \frac{\partial f}{\partial y}$:

TODO: align them vertically centered in a nice way

$$f_x = 8x - 2y + 3$$

$$f_y = -2x + 6y - 2$$

$$f_{xx} = 8 > 0$$

$$f_{yy} = 6$$

$$f_{xy} = -2$$

The critical point is determined by solving:

$$f_x = 0 \quad \text{and} \quad f_y = 0 \quad \Leftrightarrow$$

$$8x - 2y + 3 = 0 \quad \text{and} \quad -2x + 6y - 2 = 0 \quad \Leftrightarrow$$

$$x = -\frac{7}{22} \quad y = \frac{5}{22}$$

Therefore the only critical point is that $(x_0, y_0) = (-\frac{7}{22}, \frac{5}{22})$. Evaluate the Hessian at it.

$$\mathbf{H}_f(x_0, y_0) = \begin{bmatrix} 8 & -2 \\ -2 & 6 \end{bmatrix}$$

For matrix $\mathbf{H}_f(x_0, y_0)$, we have $8 > 0$ and $\det(\mathbf{H}_f)(x_0, y_0) = 44 > 0$ therefore it is positive definite. Alternatively, we could find the eigenvalues $\lambda_1 = 7 + \sqrt{5} > 0$, $\lambda_2 = 7 - \sqrt{5} > 0$ so again, according to Theorem 3.1 it is positive definite. The definitiveness combined with $f_{xx}(x_0, y_0) > 0$ guarantee that (x_0, y_0) is a local minimum and as it's the only minimum, it's also global. ■

EXAMPLE 4.6. *Find and classify the critical points of the function $f(x, y) = 3x^2y + y^3 - 3x^2 - 3y^2 + 1$ [10].*

Solution:

Find the partial derivatives: **TODO:** align them vertically centered in a nice way

$$f_x = 6xy - 6x$$

$$f_y = 3x^2 + 3y^2 - 6y$$

$$f_{xx} = 6y - 6 > 0$$

$$f_{yy} = 6y - 6$$

$$f_{xy} = 6x$$

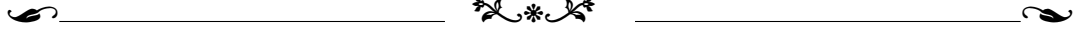
Both f_{xx}, f_{yy} are continuous. By solving $f_x = 0$, $f_y = 0$ we find that the critical points are $(0, 0), (0, 2), (1, 1), (-1, 1)$. Compute the Hessian:

$$\mathbf{H}_f = \begin{bmatrix} 6y - 6 & 6x \\ 6x & 6y - 6 \end{bmatrix}$$

For each critical point, we have:

1. At $(0, 0)$, $\det(\mathbf{H}_f) = 36 > 0$ and $f_{xx} = -6 < 0$ therefore local maximum.
2. At $(0, 2)$, $\det(\mathbf{H}_f) = 36 > 0$ and $f_{xx} = 6 > 0$ therefore local minimum.
3. At $(1, 1)$, $\det(\mathbf{H}_f) = -36 < 0$ therefore saddle point.
4. At $(-1, 1)$, $\det(\mathbf{H}_f) = -36 < 0$ therefore saddle point.

■



4.2.5 Constrained optimisation with linear constraint

We now focus on the following problem.

$$\begin{aligned} \text{Minimise } f(\mathbf{x}) &= \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0 \\ \text{s. t. } \mathbf{A} \mathbf{x} &= \mathbf{b}, \end{aligned} \tag{4.11}$$

where $\mathbf{H} = \mathbf{H}^T \in \mathbb{R}^{n \times n}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $c_0 \in \mathbb{R}$.

The set \mathcal{F} of all vectors $\mathbf{x} \in \mathbb{R}^n$ that satisfy the constraint $\mathbf{A} \mathbf{x} = \mathbf{b}$ is called feasible solution. Therefore $\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A} \mathbf{x} = \mathbf{b}\}$.

- If the system $\mathbf{A} \mathbf{x} = \mathbf{b}$ has no solution, i.e. $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$, then $\mathcal{F} = \emptyset$ therefore there is no constrained minimisation problem.
- If the system $\mathbf{A} \mathbf{x} = \mathbf{b}$ has exactly one solution, then this solution is also the unique optimal solution to the problem Eq. (4.11), since there is no other feasible solution which is better.
- The only non-trivial case is when $\mathbf{A} \mathbf{x} = \mathbf{b}$ has multiple different solution, which is the case when \mathbf{A} has more columns than rows, i.e. $m > n$ and the columns are lin. independent, therefore span \mathbb{R}^m . Then $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A}) \neq \emptyset$. Also $\mathcal{R}(\mathbf{A})$ spans $\mathbb{R}^m \Rightarrow \dim(\mathcal{R}(\mathbf{A})) = m$ and $\dim(\mathcal{N}(\mathbf{A})) = n - m$.

So for the rest of this section we assume $n > m$, which means that:

$$\mathbf{b} \in \mathcal{R}(\mathbf{A}), \quad \mathcal{N}(\mathbf{A}) \neq \emptyset \tag{4.12}$$

We also have the following lemma about a feasible solution $\bar{\mathbf{x}}$ of the system and any other solution \mathbf{x} .

LEMMA 4.5. *If $\bar{\mathbf{x}}$ is a feasible solution of $\mathbf{A} \mathbf{x} = \mathbf{b}$ then*

$$\mathbf{x} \in \mathcal{F} \iff \bar{\mathbf{x}} - \mathbf{x} \in \mathcal{N}(\mathbf{A})$$

Proof.

(\Rightarrow) :

$$\begin{aligned} \bar{\mathbf{x}} \in \mathcal{F} &\Rightarrow \mathbf{A} \bar{\mathbf{x}} = \mathbf{b} \quad \mathbf{x} \in \mathcal{F} \Rightarrow \mathbf{A} \mathbf{x} = \mathbf{b} \\ \therefore \mathbf{A}(\bar{\mathbf{x}} - \mathbf{x}) &= \mathbf{b} - \mathbf{b} = \mathbf{0} \\ \therefore \bar{\mathbf{x}} - \mathbf{x} &\in \mathcal{N}(\mathbf{A}) \end{aligned}$$

(\Leftarrow) :

If $\bar{\mathbf{x}}$ is a feasible solution and $\bar{\mathbf{x}} - \mathbf{x} \in \mathcal{N}(\mathbf{A})$, we want to show that \mathbf{x} is also a feasible solution.

$$\begin{aligned} \bar{\mathbf{x}} - \mathbf{x} \in \mathcal{N}(\mathbf{A}) &\Rightarrow \mathbf{A}(\bar{\mathbf{x}} - \mathbf{x}) = \mathbf{0} \\ \therefore \mathbf{A} \mathbf{x} &= \mathbf{A}(\bar{\mathbf{x}} - (\bar{\mathbf{x}} - \mathbf{x})) = \mathbf{A} \bar{\mathbf{x}} - \mathbf{A}(\bar{\mathbf{x}} - \mathbf{x}) = \mathbf{b} - \mathbf{0} = \mathbf{b} \end{aligned}$$

□

Let k be the dimension of $\mathcal{N}(\mathbf{A})$ and $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k$ be the coefficient vectors that are its basis.

$$\mathbf{Z} = [\mathbf{z}_1 \quad \dots \quad \mathbf{z}_k]$$

Let \mathbf{v} be the $k \times 1$ coordinate vector. Then $\bar{\mathbf{x}} - \mathbf{x} \in \mathcal{N}(\mathbf{A})$ iff $\bar{\mathbf{x}} - \mathbf{x} = \mathbf{v}\mathbf{Z}$ for some arbitrary $\mathbf{v} \in \mathbb{R}^k$ [11], [12], therefore for any feasible solution we get the representation:

$$\mathbf{x} \in \mathcal{F} \text{ iff } \mathbf{x} = \bar{\mathbf{x}} + \mathbf{Z}\mathbf{v} \text{ for some } \mathbf{v} \in \mathbb{R}^k$$

Since $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k$ are linearly independent, it follows that for every $\mathbf{x} \in \mathcal{F}$ there is a unique $\mathbf{v} \in \mathbb{R}^k$.

Therefore we want to minimise $f(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v})$. The following two lemmas show how $f(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v})$ can be expanded.

LEMMA 4.6. *If $t \in \mathbb{R}$, $\mathbf{d} \in \mathbb{R}^n$, then for the quadratic function in Eq. (4.6) we have [13]*

$$f(\mathbf{x} + t\mathbf{d}) = f(\mathbf{x}) + t(\mathbf{H}\mathbf{x} + \mathbf{c})^T \mathbf{d} + \frac{1}{2}t^2 \mathbf{d}^T \mathbf{H} \mathbf{d} \quad \forall \mathbf{x} \in \mathbb{R}^n$$

Proof.

$$\begin{aligned} f(\mathbf{x} + t\mathbf{d}) &= \frac{1}{2}(\mathbf{x} + t\mathbf{d})^T \mathbf{H}(\mathbf{x} + t\mathbf{d}) + \mathbf{c}^T(\mathbf{x} + t\mathbf{d}) + c_0 \\ &= \frac{1}{2}(\mathbf{x}^T \mathbf{H} \mathbf{x} + t\mathbf{d}^T \mathbf{H} \mathbf{x} + t\mathbf{x}^T \mathbf{H} \mathbf{d} + t^2 \mathbf{d}^T \mathbf{H} \mathbf{d}) + \mathbf{c}^T \mathbf{x} + t\mathbf{c}^T \mathbf{d} + c_0 \\ &= \left(\frac{1}{2}\mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0\right) + t\left(\frac{1}{2}\mathbf{d}^T \mathbf{H} \mathbf{x} + \frac{1}{2}\mathbf{x}^T \mathbf{H} \mathbf{d} + \mathbf{c}^T \mathbf{d}\right) + \frac{1}{2}t^2 \mathbf{d}^T \mathbf{H} \mathbf{d} \\ &= f(\mathbf{x}) + t(\mathbf{H}\mathbf{x} + \mathbf{c})^T \mathbf{d} + \frac{1}{2}t^2 \mathbf{d}^T \mathbf{H} \mathbf{d} \end{aligned}$$

□

LEMMA 4.7. *From Lemma 4.6 for $t = 1$, $\mathbf{d} = \mathbf{y} - \mathbf{x}$ we obtain [13]*

$$f(\mathbf{y}) = f(\mathbf{x}) + (\mathbf{H}\mathbf{x} + \mathbf{c})^T \mathbf{d} + \frac{1}{2}t^2 \mathbf{d}^T \mathbf{H} \mathbf{d}$$

Now we can compute $f(\bar{\mathbf{x}} + \mathbf{Z}\mathbf{v})$ using Lemma 4.7 for $d = 1$, $\mathbf{d} = \mathbf{Z}\mathbf{v}$:

$$\begin{aligned} f(\bar{\mathbf{x}}) + \mathbf{Z}\mathbf{v} &= f(\bar{\mathbf{x}}) + (\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})^T \mathbf{Z}\mathbf{v} + \frac{1}{2}(\mathbf{Z}\mathbf{v})^T \mathbf{H}(\mathbf{Z}\mathbf{v}) \\ &= f(\bar{\mathbf{x}}) + (\mathbf{Z}^T(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}))^T \mathbf{v} + \frac{1}{2}\mathbf{v}^T (\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \mathbf{v} \end{aligned}$$

Therefore we have transformed the problem to an unconstrained one, namely

$$\begin{aligned} &\text{minimise } f(\bar{\mathbf{x}}) + (\mathbf{Z}^T(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}))^T \mathbf{v} + \frac{1}{2}\mathbf{v}^T (\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \mathbf{v} \\ &\text{s.t. } \mathbf{v} \in \mathbb{R}^k \end{aligned} \tag{4.13}$$

f does have a minimum if it convex. Without proving it, is it convex when.

$$\mathbf{Z}^T \mathbf{H} \mathbf{Z} \text{ is positive semi-definite} \tag{4.14}$$

The following condition must also hold

$$(\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \hat{\mathbf{v}} = -\mathbf{Z}^T(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$$

Therefore $\hat{\mathbf{x}} \in \mathcal{F}$ is an optimal solution to Eq. (4.13) iff

$$(\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \hat{\mathbf{v}} = -\mathbf{Z}^T(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c}), \quad \hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}}, \quad \text{with } \mathbf{A}\bar{\mathbf{x}} = \mathbf{b} \tag{4.15}$$

In case $(\mathbf{Z}^T \mathbf{H} \mathbf{Z})$ is positive definite, then is it invertible too and the system $(\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \hat{\mathbf{v}} = -\mathbf{Z}^T(\mathbf{H}\bar{\mathbf{x}} + \mathbf{c})$ has a unique solution $\hat{\mathbf{v}}$ and then $\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z}\hat{\mathbf{v}}$ is the unique optimal solution to problem Eq. (4.13). To reiterate, we have proven the following theorem.

THEOREM 4.6. $\hat{\mathbf{x}}$ is optimal solution to Eq. (4.11) if:

1. $\mathbf{Z}^T \mathbf{H} \mathbf{Z}$ is positive semi-definite
2. $\hat{\mathbf{x}} = \mathbf{Z}^T + \mathbf{Z} \hat{\mathbf{v}}$, where $(\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \hat{\mathbf{v}} = -\mathbf{Z}^T (\mathbf{H} \bar{\mathbf{x}} + \mathbf{c})$, $\bar{\mathbf{x}} \in \mathcal{F}$

$\bar{\mathbf{H}} = \mathbf{Z}^T \mathbf{H} \mathbf{Z}$ is also called reduced Hessian.

EXAMPLE 4.7. [7] Let $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x} + c_0$, where

$$\mathbf{H} = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad c_0 = 0.$$

Consider the minimisation of f over the feasible region $\mathcal{F} = \{x | \mathbf{A} \mathbf{x} = \mathbf{b}\}$, where

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \end{bmatrix}, \quad b = 2$$

, i.e. the minimisation problem

$$\min -x_1^2 + \frac{1}{2}x_2^2 + x_1 + x_2 \text{ s.t. } 2x_1 + x_2 = 2$$

Solution:

Find the null space \mathbf{Z} .

$$\mathbf{A} \mathbf{x} = \mathbf{0} \Leftrightarrow$$

$$2x_1 + x_2 = 0$$

Setting $s := x_2$ as the free variable, we have

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} -1/2 \\ 1 \end{bmatrix}}_{\mathbf{Z}} s$$

Therefore the matrix that spans the null space is $[-1/2 \quad 1]^T$ or equivalently and for convenience $\mathbf{Z} = [-1 \quad 2]^T$. The reduced Hessian is given by

$$\bar{\mathbf{H}} = \mathbf{Z}^T \mathbf{H} \mathbf{Z} = 2 > 0$$

Therefore it is positive semi-definite hence the problem has a unique solution. To find a solution, start with an $\bar{\mathbf{x}} \in \mathcal{F}$. We can take, for example, $\bar{\mathbf{x}} = [1 \quad 0]^T$.

Solving $(\mathbf{Z}^T \mathbf{H} \mathbf{Z}) \hat{\mathbf{v}} = -\mathbf{Z}^T (\mathbf{H} \bar{\mathbf{x}} + \mathbf{c})$, $\hat{\mathbf{v}}$ can be found.

$$\bar{\mathbf{H}} \hat{\mathbf{v}} = -\mathbf{Z}^T (\mathbf{H} \bar{\mathbf{x}} + \mathbf{c}) \Leftrightarrow$$

$$2\hat{\mathbf{v}} = -\mathbf{Z} = -3 \Leftrightarrow$$

$$\hat{\mathbf{v}} = -\frac{3}{2}$$

Then

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{Z} \hat{\mathbf{v}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} - \frac{3}{2} = \begin{bmatrix} \frac{5}{2} \\ -3 \end{bmatrix}$$

4.2.6 General constrained optimisation

TODO: add it in the future, see comment

5 Singular value decomposition (SVD)

5.1 Introduction

Singular Value Decomposition (SVD) allows any $m \times n$ matrix (not just symmetric) to be factorised as a product of three matrices. We will derive SVD in a number of steps [14].

1. For an arbitrary $m \times n$ matrix \mathbf{A} , matrices $\mathbf{A}^T \mathbf{A}$ and $\mathbf{A} \mathbf{A}^T$ are symmetric.
2. Therefore $\mathbf{A}^T \mathbf{A}$ has r orthonormal eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ pertaining to eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_r$. r is the rank of $\mathbf{A}^T \mathbf{A}$. We used that the rank of a matrix is equal to its number of non-zero eigenvectors (proof in Appendix A.7). Therefore $\mathbf{A}^T \mathbf{A}$ can be orthogonally diagonalised.
3. Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be the orthonormal basis for \mathbb{R}^n consisting of the eigenvectors of $\mathbf{A}^T \mathbf{A}$. Then for $1 \leq i \leq n$:

$$\begin{aligned}\|\mathbf{A}\mathbf{v}_i\|^2 &= (\mathbf{A}\mathbf{v}_i)^T \mathbf{A}\mathbf{v}_i = \mathbf{v}_i^T (\mathbf{A}^T \mathbf{A} \mathbf{v}_i) \\ &= \mathbf{v}_i^T \lambda_i \mathbf{v}_i = \lambda_i \|\mathbf{v}_i\|^2 \implies \\ \|\mathbf{A}\mathbf{v}_i\| &= \sqrt{\lambda_i}\end{aligned}$$

So the eigenvalues of $\mathbf{A}^T \mathbf{A}$ are non-negative and we assume the ordering $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$.

We define

$$\sigma_i = \sqrt{\lambda_{\mathbf{A}^T \mathbf{A} \ i}} \quad (5.1)$$

the singular values of \mathbf{A} . The singular value σ_i is equal to the length of the vector $\mathbf{A}\mathbf{v}_i$. Don't forget that λ_i are the eigenvalues of $\mathbf{A}^T \mathbf{A}$.

4. Define

$$\mathbf{u}_i = \frac{\mathbf{A}\mathbf{v}_i}{\|\mathbf{A}\mathbf{v}_i\|} = \frac{\mathbf{A}\mathbf{v}_i}{\sigma_i} \quad (5.2)$$

☞ **Comment:** [15] We want \mathbf{U} to be square $n \times n$ and orthogonal to allow the decomposition. When using Eq. (5.2) we may run into a problem where we have used all of the non-zero singular values to find the columns that make up \mathbf{U} , but do not have a square orthogonal matrix. In this case, use the Gram-Schmidt process to find a basis for the subspace spanned by the existing columns of \mathbf{U} . The strategy is to find a random vector \mathbf{y} in the plane spanned by \mathbf{U} , e.g. if \mathbf{U} is 3×2 we want another 3×1 vector to make it orthonormal. Then we can plug it into the Gram-Schmidt projection formula to find the required remaining column vectors of \mathbf{U} .

$$\mathbf{u} = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 - \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 - \dots - \frac{\mathbf{y} \cdot \mathbf{u}_a}{\mathbf{u}_a \cdot \mathbf{u}_a} \mathbf{u}_a \quad (5.3)$$

, where a is the number of eigenvectors we had already found. Any column vector \mathbf{u} found must be converted to unit vector so it can be placed in matrix \mathbf{U} .

It is easy to show that $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ is an orthonormal basis for \mathbb{R}^r , $1 \leq i \leq r$. Writing all generated equations for all i in matrix form we have:

$$\begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_r^T \end{bmatrix} \mathbf{A} [\mathbf{v}_1 \ \dots \ \mathbf{v}_r] = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

5. If $r < m$, then we can use Gram-Schmidt as described in the comment so we can pick unit vectors $\mathbf{u}_{r+1}, \dots, \mathbf{u}_m$ such that the basis $\mathbf{u}_1, \dots, \mathbf{u}_m$ is orthonormal. We then have that $\mathbf{u}_i \mathbf{A} \mathbf{v}_i = \sigma_i$, $1 \leq i \leq r$ and $\mathbf{u}_i \mathbf{A} \mathbf{v}_i = 0$, $r+1 \leq i \leq m$.
6. Therefore for all cases we have

$$\begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix} \mathbf{A} [\mathbf{v}_1 \ \dots \ \mathbf{v}_n] = \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \vdots \\ & & \sigma_r & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix}$$

To conclude, if we let $\mathbf{U} = [\mathbf{u}_1 \ \dots \ \mathbf{u}_m]$, $\mathbf{V} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$, and Σ the diagonal matrix of singular values padded with zeros, we can write

$$\mathbf{U}^T \mathbf{A} \mathbf{V} = \Sigma$$

, and because both \mathbf{U}, \mathbf{V} are orthogonal, i.e. $\mathbf{U}^T \mathbf{U} = \mathbf{I}, \mathbf{V}^T \mathbf{V} = \mathbf{I}$ we can solve for \mathbf{A} as follows

$$\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^T \quad (5.4)$$

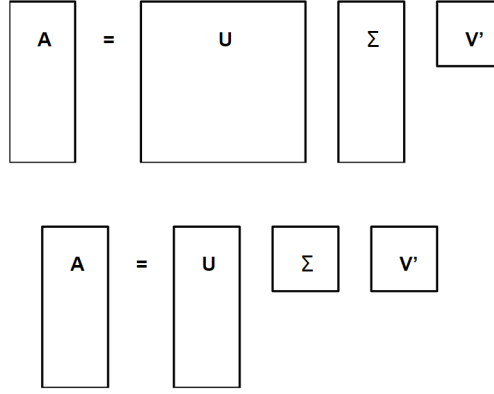


Figure 9: SVD component matrices visualised, full (above) and "economy", i.e. truncated to rank r (below).



To reiterate, we have proven the Singular Value Decomposition of \mathbf{A} . Remember that \mathbf{V} is the eigenvector matrix of $\mathbf{A}^T \mathbf{A}$, Σ the pseudo-diagonal matrix of singular values, and \mathbf{U} the unit column space of \mathbf{A} , given by $\mathbf{A}\mathbf{v}_i/\sigma_i$.

An alternative and sometimes easier way to find \mathbf{U} , \mathbf{V} is described below.

DEFINITION 5.1. Matrix \mathbf{U} in Eq. (5.4) is called left singular matrix and matrix \mathbf{V} is called right singular matrix.

LEMMA 5.1. Matrix \mathbf{U} in Eq. (5.4) (left singular) contains the unit eigenvectors of $\mathbf{A}\mathbf{A}^T$ as its columns and matrix \mathbf{V} contains the unit eigenvectors of $\mathbf{A}^T \mathbf{A}$ as its columns.

Proof.

We know from the SVD that:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$$

, where \mathbf{U} , \mathbf{V} are orthonormal and $\Sigma^T \Sigma$ (or equivalently $\Sigma \Sigma^T$) contains the eigenvalues of $\mathbf{A}^T \mathbf{A}$ (or $\mathbf{A}\mathbf{A}^T$) in its diagonal, padded with enough zeros such that its size is same \mathbf{A} 's. Therefore

$$\mathbf{A}^T = \mathbf{V}\Sigma^T \mathbf{U}^T$$

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\Sigma\mathbf{V}^T\mathbf{V}\Sigma^T\mathbf{U}^T = \mathbf{U}\Sigma\Sigma^T\mathbf{U}^T \quad (5.5)$$

$$\mathbf{A}^T \mathbf{A} = \mathbf{V}\Sigma^T \mathbf{U}^T \mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^T \Sigma \mathbf{V}^T \quad (5.6)$$

It is clear from (1) that \mathbf{U} diagonalises $\mathbf{A}\mathbf{A}^T$ and from (2) that \mathbf{V} diagonalises $\mathbf{A}^T \mathbf{A}$, which concludes the proof. \square

THEOREM 5.1. [1] Let \mathbf{A} be an $m \times n$ matrix with rank r . Then there exists an $m \times n$ matrix Σ that contains the r singular values of \mathbf{A} ($\sigma_1, \dots, \sigma_r$) in its diagonal padded with $m - r$ rows and $n - r$ columns of zeros

$$\Sigma = \left[\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ \hline & & & \mathbf{0}_{r \times n-r} & & \\ \mathbf{0}_{m-r \times r} & & & & \mathbf{0}_{m-r \times n-r} & \end{array} \right] \quad (5.7)$$

and there exist orthogonal matrices $\mathbf{U}_{m \times m}$, $\mathbf{V}_{n \times n}$ such that

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T, \quad (5.8)$$

☞ **Comment:** The SVD exists for any matrix \mathbf{A} . There are no size or form restrictions.

EXAMPLE 5.1. Find the singular values of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

. Solution:

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

Find its eigenvalues.

$$\det\left(\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} - \lambda \mathbf{I}\right) = 0 \Rightarrow$$

$$\lambda^2 - 10\lambda + 9 = 0 \Rightarrow$$

$$\lambda = 9 \text{ or } \lambda = 1$$

The singular values of \mathbf{A} are $\sigma_1 = \sqrt{9}$, $\sigma_2 = \sqrt{1}$.

EXAMPLE 5.2. Find the Singular Value Decomposition of matrix $\mathbf{A} = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$. [1]

Solution:

Find $\mathbf{A}^T \mathbf{A}$ and its eigen-data.

$$\mathbf{A}^T \mathbf{A} = 10 \begin{bmatrix} 8 & 10 & 4 \\ 10 & 17 & 14 \\ 4 & 14 & 20 \end{bmatrix}$$

The eigenvalues in decreasing order are $\lambda_1 = 360, \lambda_2 = 90, \lambda_3 = 0$ therefore the singular values are $\sigma_1 = \sqrt{360}, \sigma_2 = \sqrt{90}, \sigma_3 = 0$. If we created a diagonal matrix with the non-zero eigenvalues it would be 2×2 . We pad with zeros to make it 2×3 .

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix}$$

The corresponding eigenvectors are

$$\mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

$\mathbf{A}\mathbf{A}^T$ is small (2×2) so we could find its unit eigenvectors from Eq. (5.5) to construct \mathbf{U} .

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 333 & 81 \\ 81 & 117 \end{bmatrix}, \quad \lambda_1 = 360 : \mathbf{u}_1 = \begin{bmatrix} 0.948 \\ 0.316 \end{bmatrix}, \quad \lambda_2 = 90 : \mathbf{u}_2 = \begin{bmatrix} 0.316 \\ -0.948 \end{bmatrix}$$

$$\therefore \mathbf{U} = \begin{bmatrix} 0.948 & 0.316 \\ 0.316 & -0.948 \end{bmatrix}$$

An alternative way to construct \mathbf{U} would be using Eq. (5.2). Verify that they are the same.

$$\mathbf{u}_1 = \frac{\mathbf{A}\mathbf{v}_1}{\sigma_1} = \dots$$

$$\mathbf{u}_2 = \frac{\mathbf{A}\mathbf{v}_2}{\sigma_2} = \dots$$

Finally, we have the SVD decomposition

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \begin{bmatrix} 0.948 & 0.316 \\ 0.316 & -0.948 \end{bmatrix} \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

The last equation is easy to numerically verify e.g. with Matlab.

EXAMPLE 5.3. Compute the SVD of the matrix $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ [15].

Solution:

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Eigenvalues and unit eigenvectors:

$$\lambda_1 : \mathbf{v}_1 = \begin{bmatrix} 1\sqrt{2} \\ 1\sqrt{2} \end{bmatrix}, \quad \lambda_2 : \mathbf{v}_2 = \begin{bmatrix} -1\sqrt{2} \\ 1\sqrt{2} \end{bmatrix}$$

The unit eigenvectors of $\mathbf{A}^T \mathbf{A}$ are the columns of the right singular matrix \mathbf{V} :

$$\mathbf{V} = \begin{bmatrix} 1\sqrt{2} & -1\sqrt{2} \\ 1\sqrt{2} & 1\sqrt{2} \end{bmatrix}$$

The square roots of the eigenvalues of $\mathbf{A}^T \mathbf{A}$ are the diagonal entries of $\mathbf{\Sigma}$:

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Finally, use Eq. (5.2) to compute the left singular matrix \mathbf{U} .

$$\mathbf{u}_1 = \frac{\mathbf{A}\mathbf{v}_1}{\sigma_1} = \frac{1}{\sqrt{3}} \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right) = \begin{bmatrix} \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{\mathbf{A}\mathbf{v}_2}{\sigma_2} = \frac{1}{1} \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \right) = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Therefore the \mathbf{U} matrix is

$$\mathbf{U} = \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

We have run into the problem described in the comment in the theory - \mathbf{U} is not square therefore the SVD multiplication is impossible. We want to make it a 3×3 orthonormal matrix. The way to do this is by picking a random vector (a good idea is one with a lot of zeros) in the plane spanned by the columns of \mathbf{U} , i.e. in \mathbb{R}^3 . We pick

$$\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Use the Gram-Schmid formula in Eq. (5.3) to construct an orthonormal basis $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ in \mathbb{R}^3 starting from $\mathbf{u}_1, \mathbf{u}_2$.

$$\mathbf{u}_3 = \mathbf{y} - \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \mathbf{u}_1 - \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \mathbf{u}_2 = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{bmatrix},$$

$$\frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{bmatrix}$$

Therefore we have computed the left singular matrix and everything required for the SVD.

$$\mathbf{U} = \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix}$$

To sum up:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \Leftrightarrow$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1\sqrt{2} & -1\sqrt{2} \\ 1\sqrt{2} & 1\sqrt{2} \end{bmatrix}^T$$

■

☞ **Comment:** We saw in the last example that the SVD had the form

$$\mathbf{A} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3] \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix}$$

And expanding the matrix product for \mathbf{A} into a sum:

$$\begin{aligned} \mathbf{A} &= \sqrt{3} \begin{bmatrix} \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + 1 \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \sqrt{3} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} \end{bmatrix}}_{\mathbf{A}_1} + \underbrace{\begin{bmatrix} 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{\mathbf{A}_2} \end{aligned}$$

So \mathbf{A} is written as a linear combination of two rank-one matrices. Their rank is not a coincidence and why it's 1 is explained below. The first term gives a better approximation to \mathbf{A} than the second as its weight (σ_1) is higher. Remember that singular values are in non-decreasing order. Let $\tilde{\mathbf{A}}_1, \tilde{\mathbf{A}}_2$ be the approximations of \mathbf{A} by the first and second term respectively, then we can see \mathbf{A}_1 is closer to the original matrix.

$$\tilde{\mathbf{A}}_1 = \begin{bmatrix} 1 & 1 \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \tilde{\mathbf{A}}_2 = \begin{bmatrix} 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Approximating a matrix by using the product of the first k values, left and right singular vectors and ignoring the remaining products is called low rank approximation.

DEFINITION 5.2. If $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$ then the product $\mathbf{u}\mathbf{v}^T$ is called outer product. It yields an $m \times n$ matrix whose rank is 1.

Proof.

Let the product be \mathbf{A} .

$$\mathbf{A} = \mathbf{u}\mathbf{v}^T = \begin{bmatrix} u_1\mathbf{v}^T \\ \vdots \\ u_m\mathbf{v}^T \end{bmatrix}$$

So the rows are all linearly dependent and by Gaussian elimination we can only leave $u_1\mathbf{v}^T$, i.e. $\text{rank}(\mathbf{A}) = 1$ [16]. \square

Proof.

Take a vector $\mathbf{x} \in \mathbb{R}^n$. Then the outer product matrix maps it to another vector in \mathbb{R}^n .

$$\mathbf{A}\mathbf{x} = (\mathbf{u}\mathbf{v}^T)\mathbf{x} = \mathbf{u}(\mathbf{v}^T\mathbf{x}) = k\mathbf{u}$$

$k = \mathbf{v}^T\mathbf{x}$ is a scalar, particularly it is a dot product $((1 \times n) \times (n \times 1) \rightarrow 1)$. Hence $\text{rank}(\mathbf{A}) = \dim(\mathcal{R}(\mathbf{A})) = 1$ [17]. \square

Outer products are useful as they appear in the expanded SVD before the low rank approximation we will discuss the geometry of SVD.

5.1.1 Geometric interpretation of SVD

We reiterate SVD's equation.

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

We will consider the product

$$\mathbf{A}\mathbf{x} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{x}, \quad \|\mathbf{x}\| = 1$$

We know that orthogonal matrices represent rotations (Appendix A.8. If \mathbf{A} is $m \times n$, then \mathbf{U} is $m \times m$, **Sigma** is $m \times n$ and \mathbf{V} is $n \times n$. As we will see, SVD specifies that every linear transformation is fundamentally a rotation or reflection, followed by a scaling, followed by another rotation or reflection [18].

1. A rotation of vectors in the n -dimensional domain represented by matrix \mathbf{V}
2. A linear dilation or contraction (essentially scalar multiplication) along each of the m coordinate directions represented by $\mathbf{\Sigma}$.
3. A rotation of vectors in the m -dimensional range represented by \mathbf{U} .

Below is a visualisation of those actions for a 2×2 matrix operator \mathbf{A} on unit column vector $[x_1 \ x_2]^T$. In the illustration both rotations are in \mathbb{R}^2 but they could be in different dimensions.

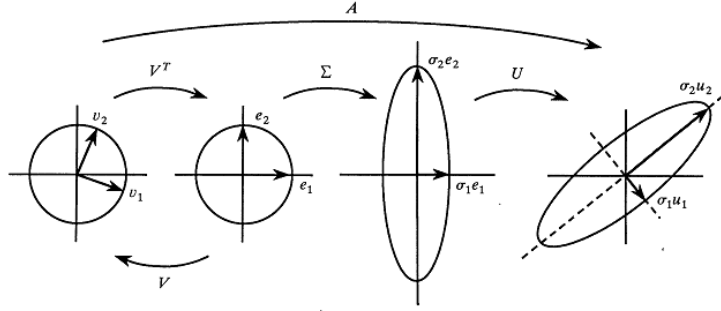


Figure 10: Linear transform from \mathbb{R}^2 to \mathbb{R}^2 [19].

We also know from Section 4, Eq. (4.2), that the maximum value of $\|\mathbf{Ax}\|$ is $\sqrt{\lambda_1} = \sigma_1$ and the minimum $\sqrt{\lambda_1} = \sigma_1$, $\lambda_1 \geq \lambda_2 \geq 0$. Therefore SVD transform the unit circle in this case (or "unit sphere" in general).

EXAMPLE 5.4. If $\mathbf{A} = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ then the transform $\mathbf{x} \rightarrow \mathbf{Ax}$ maps the unit sphere $\mathbf{x} : \|\mathbf{x}\| = 1$ in \mathbb{R}^3 onto an ellipse in \mathbb{R}^2 , as shown in Fig. 10. Find the unit vectors for which $\|\mathbf{Ax}\|$ and compute the minimum and maximum lengths.

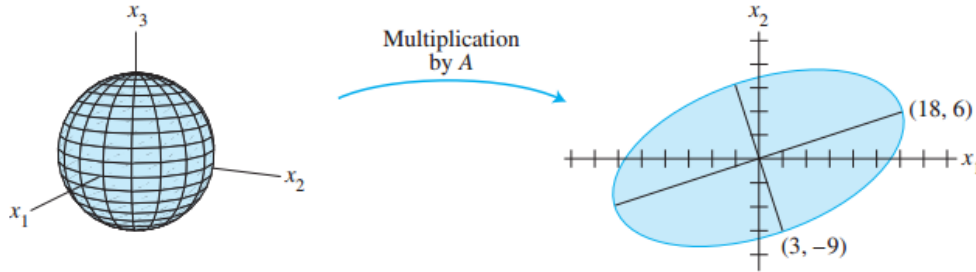


Figure 11: SVD consists of 3 action; rotation in one domain, followed by scaling, followed by rotation in another domain [1].

Solution:

From Section 4, $\|\mathbf{Ax}\|$ is maximised (minimised) for the eigenvector corresponding to the maximum (minimum) non-zero eigenvalue λ_1 (λ_i , $i > 1$) of $\mathbf{A}^T \mathbf{A}$.

$$\mathbf{A}^T \mathbf{A} = 10 \begin{bmatrix} 8 & 10 & 4 \\ 10 & 17 & 14 \\ 4 & 14 & 20 \end{bmatrix}$$

Eigen-data:

$$\lambda_1 = 360 : \mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \lambda_2 = 90 : \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \lambda_3 = 0 : \mathbf{v}_3 = \begin{bmatrix} -2/3 \\ 2/3 \\ -1/3 \end{bmatrix}$$

We then find that $\max_{\|\mathbf{x}\|=1} (\mathbf{Ax}) = 6\sqrt{10}$, $\min_{\|\mathbf{x}\|=1} (\mathbf{Ax}) = \sqrt{90}$.

We will find how the transform $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$, $\|\mathbf{x}\| = 1$ maps vector \mathbf{x} to. From SVD we know:

$$\mathbf{A}\mathbf{V} = \mathbf{\Sigma}\mathbf{U}$$

We know that

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} 18 \\ 6 \end{bmatrix}, \quad \mathbf{A}\mathbf{v}_2 = \begin{bmatrix} 3 \\ -9 \end{bmatrix}, \quad \mathbf{A}\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So we write the coordinate vector \mathbf{x} in terms of the basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$:

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3$$

A general point given by the vector $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3$ is mapped to

$$\mathbf{A}\mathbf{V}\mathbf{x} = \begin{bmatrix} 18x_1 + 3x_2 \\ 6x_1 - 9x_2 \end{bmatrix}$$

Because \mathbf{x} is on the unit sphere, we can parametrise it as $(x_1, x_2, x_3) = (r \cos \theta, r \sin \theta, z)$, so the polar coordinates w.r.t. $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are

$$\begin{aligned} \mathbf{A}\mathbf{V}\mathbf{x} &= \begin{bmatrix} r18 \cos \theta + r3 \sin \theta \\ r6 \cos \theta - r9 \sin \theta \end{bmatrix} \\ &= \begin{bmatrix} r6\sqrt{10}\cos(\alpha - \theta) \\ r3\sqrt{10}\sin(\alpha - \theta) \end{bmatrix} \end{aligned}$$

In vector notation, we have the mapping $(\sin(\theta), \cos(\theta), z) \mapsto (r6\sqrt{10}\cos(\alpha - \theta), r3\sqrt{10}\sin(\alpha - \theta))$, where $r = 1$ because \mathbf{x} is on the unit sphere. For example, point $(0, 0, 1)$ maps to $(0, 0)$ [20]. ■

COROLLARY 5.1.

- As any matrix \mathbf{A} has an SVD decomposition, it will always transform the "unit sphere" (hypersphere) into an ellipsoid.
- If \mathbf{A} does not have full rank, then some of the semi-axes of the transformed ellipsoid will have length 0.

To formalise what was observed in the last example, if we have a matrix $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, a coordinates vector on the unit sphere in \mathbb{R}^n and express it w.r.t. right singular vectors \mathbf{V} , i.e.

$$\left\{ \mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_n\mathbf{v}_n \mid \sum_{i=1}^n x_i^2 \right\}$$

Then from SVD we know that the image of \mathbf{x} is

$$\mathbf{A}\mathbf{x} = \sum_{i=1}^{r=\text{rank}(\mathbf{A})} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \mathbf{v}_i x_i = \sum_{i=1}^{r=\text{rank}(\mathbf{A})} \sigma_i \mathbf{u}_i x_i$$

Letting $y_i = \sigma_i x_i$, we see that the image of the unit sphere consists of the vectors $\mathbf{y}_1\mathbf{u}_1 + \mathbf{y}_2\mathbf{u}_2 + \dots + \mathbf{y}_r\mathbf{u}_r$, where [21]

$$\begin{aligned} \frac{y_1^2}{\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} + \dots + \frac{y_r^2}{\sigma_r^2} &= \sum_{i=1}^r x_i^2 \\ &\leq 1 \end{aligned}$$

If \mathbf{A} has full column rank, so $r = n$ the inequality is actually a strict equality. Otherwise, some of the x_i are missing on the right, and the sum can be anything from 0 to 1. This leads to the next corollary.

COROLLARY 5.2. \mathbf{A} maps the unit sphere of \mathbb{R}^n to a k -dimensional ellipsoid with semi-axes in the directions \mathbf{u}_i and with the magnitudes σ_i . If $r = n$ the image is just the surface of the ellipsoid, otherwise it is the solid ellipsoid. In summary, we can visualize the effect of \mathbf{A} follows:

It first collapse $n - r$ dimensions of the domain, then distorts the remaining dimensions, stretching and squeezing the unit r -sphere into an ellipsoid, and finally embeds the ellipsoid in \mathbb{R}^m .

5.2 Application 1: SVD and fundamental subspaces

Motivated by the Σ matrix partitioning in the SVD Theorem 5.1, we can partition \mathbf{U} and \mathbf{V} in a similar way. Let \mathbf{A} be $m \times n$, then, denoting $\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r)$ [22].

$$\begin{aligned}
 \mathbf{A} &= \left[\underbrace{\mathbf{U}_1}_{m \times r} \underbrace{\mathbf{U}_2}_{m \times m-r} \right] \left[\begin{array}{c|c} \Sigma_r & \mathbf{0}_{r \times n-r} \\ \hline \mathbf{0}_{m-r \times r} & \mathbf{0}_{m-r \times n-r} \end{array} \right] \left[\underbrace{\mathbf{V}_1}_{n \times r} \underbrace{\mathbf{V}_2}_{n \times n-r} \right]^T \\
 &= \left[\underbrace{\mathbf{U}_1 \Sigma_r}_{m \times r} \underbrace{\mathbf{0}}_{m \times n-r} \right] \left[\begin{array}{c} \mathbf{V}_1^T \\ \hline \mathbf{V}_2^T \end{array} \right] \\
 &= \underbrace{\mathbf{U}_1}_{m \times r} \underbrace{\Sigma_r}_{r \times r} \underbrace{\mathbf{V}_1^T}_{r \times n} \quad (5.9)
 \end{aligned}$$

Eq. (5.9) is a more compact form of SVD for full rank.

COROLLARY 5.3. Suppose \mathbf{A} is any $n \times n$ matrix, and $\mathbf{A} = \mathbf{U}_1 \Sigma_r \mathbf{V}_1$ is the full rank SVD. It follows that an orthonormal set of basis vectors for the range (or column space) $\mathcal{R}(\mathbf{A})$ of \mathbf{A} are the columns of \mathbf{U}_1 [22].

Proof. When $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ (in the column space) then for some vector \mathbf{x} :

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \mathbf{U}_1 \Sigma_r \mathbf{V}_1^T \mathbf{x} = \mathbf{U}_1 \mathbf{x}^*$$

, where $\mathbf{x}^* = \Sigma_r \mathbf{V}_1^T \mathbf{x}$ is an $r \times 1$ vector. Therefore also $\mathbf{b} \in \mathcal{R}(\mathbf{U}_1)$, which means

$$\mathcal{R}(\mathbf{A}) \subseteq \mathcal{R}(\mathbf{U}_1) \quad (1)$$

Since Σ_r is invertible and symmetric and \mathbf{V}_1 is orthonormal:

$$\mathbf{U}_1 \Sigma_r \mathbf{V}_1^T = \mathbf{A} \implies \mathbf{U}_1 = \mathbf{A} \mathbf{V}_1 \Sigma_r^{-1}$$

If $\mathbf{b} \in \mathcal{R}(\mathbf{U}_1)$, then for some vector \mathbf{y} :

$$\mathbf{b} = \mathbf{U}_1 \mathbf{y} = \mathbf{A} \mathbf{V}_1 \Sigma_r^{-1} \mathbf{y} = \mathbf{A} \mathbf{y}^*$$

Therefore also $\mathbf{y} \in \mathcal{R}(\mathbf{A})$, which means

$$\mathcal{R}(\mathbf{A}) \supseteq \mathcal{R}(\mathbf{U}_1) \quad (2)$$

From Eq. (1), Eq. (2) we have $\mathcal{R}(\mathbf{U}_1) = \mathcal{R}(\mathbf{A})$. Since \mathbf{U}_1 is an orthonormal set, column vectors in \mathbf{U}_1 are an orthonormal basis for $\mathcal{R}(\mathbf{A})$. \square

In a similar fashion, we can prove the following.

COROLLARY 5.4. Suppose \mathbf{A} is any $n \times n$ matrix, and $\mathbf{A} = \mathbf{U}_1 \Sigma_r \mathbf{V}_1$ is the full rank SVD. It follows that an orthonormal set of basis vectors for the row space $\mathcal{R}(\mathbf{A}^T)$ of \mathbf{A} are the columns of \mathbf{V}_1 [22].

For the null space of \mathbf{B} we have.

COROLLARY 5.5. Suppose \mathbf{A} is any $m \times n$ rank r matrix, and $\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^T$ is the SVD. Partition \mathbf{U} into $[\mathbf{U}_1 \mathbf{U}_2]$, where \mathbf{U}_1 is the first r columns of \mathbf{U} , \mathbf{U}_2 is the last $m-r$ columns of \mathbf{U} . Similarly, partition \mathbf{V} as $[\mathbf{V}_1 \mathbf{V}_2]$. It follows that an orthonormal set of basis vectors for $N(\mathbf{A})$ are the columns of \mathbf{V}_2 .

Proof. Solving Eq. (5.9) for Σ :

$$\Sigma = \left[\begin{array}{c|c} \Sigma_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] = \left[\begin{array}{c} \mathbf{U}_1^T \\ \hline \mathbf{U}_2^T \end{array} \right] \mathbf{A} \left[\begin{array}{c|c} \mathbf{V}_1 & \mathbf{V}_2 \end{array} \right] = \left[\begin{array}{c|c} \mathbf{U}_1^T \mathbf{A} \mathbf{V}_1 & \mathbf{U}_1^T \mathbf{A} \mathbf{V}_2 \\ \hline \mathbf{U}_2^T \mathbf{A} \mathbf{V}_1 & \mathbf{U}_2^T \mathbf{A} \mathbf{V}_2 \end{array} \right] \quad (3)$$

The partitioned sub-matrices between Σ_r must match each other (do the multiplications and find the dimensions of $\mathbf{U}_i^T \mathbf{A} \mathbf{V}_i$) therefore we must have:

$$\left[\begin{array}{c} \mathbf{U}_1^T \mathbf{A} \mathbf{V}_2 \\ \hline \mathbf{U}_2^T \mathbf{A} \mathbf{V}_2 \end{array} \right] = \mathbf{0} \implies \left[\begin{array}{c} \mathbf{U}_1^T \\ \hline \mathbf{U}_2^T \end{array} \right] \mathbf{A} \mathbf{V}_2 = \mathbf{0} \implies \mathbf{U}^T \mathbf{A} \mathbf{V}_2 = \mathbf{0} \implies \mathbf{A} \mathbf{V}_2 = \mathbf{0}$$

Therefore every column vector $\mathbf{v}_{2,i}$ of \mathbf{V}_2 is in the null space of \mathbf{A} and also the columns of \mathbf{V}_2 form an orthonormal basis. Finally, we know $\dim(N(\mathbf{A})) = n - r = \dim(\mathbf{V}_2)$, so it follows that the columns of \mathbf{V}_2 form a basis for $N(\mathbf{A})$. \square

A similar result can be derived for the left null space, or the null space of \mathbf{A}^T , i.e. for all vectors such that $\mathbf{A}^T \mathbf{v} = 0$

COROLLARY 5.6. Suppose \mathbf{A} is any $m \times n$ rank r matrix, and $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is the SVD. Partition \mathbf{U} into $[\mathbf{U}_1 \ \mathbf{U}_2]$, where \mathbf{U}_1 is the first r columns of \mathbf{U} , \mathbf{U}_2 is the last $m - r$ columns of \mathbf{U} . Similarly, partition \mathbf{V} as $[\mathbf{V}_1 \ \mathbf{V}_2]$, where \mathbf{V}_1 is the first r columns of \mathbf{V} , \mathbf{V}_2 is the last $n - r$ columns of \mathbf{V} . It follows that an orthonormal set of basis vectors for the left null space of \mathbf{A} $\mathcal{N}(\mathbf{A}^T)$ are the columns of \mathbf{U}_2 [22].

Proof. Work similarly to the null space proof. Setting $[\mathbf{U}_2^T \mathbf{A} \mathbf{V}_1 \ \mathbf{U}_2^T \mathbf{A} \mathbf{V}_2] = [\mathbf{0} \ \mathbf{0}]$ we conclude that $\mathbf{A}^T \mathbf{U}_2 = \mathbf{0}$. Therefore every column vector $\mathbf{u}_{2,i}$ of \mathbf{U}_2 is in the left null space of \mathbf{A} and also the columns of \mathbf{U}_2 form an orthonormal basis. We know $\dim(\mathcal{N}(\mathbf{A}^T)) = m - r = \dim(\mathbf{U}_2)$, so it follows that the columns of \mathbf{U}_2 form a basis for $\mathcal{N}(\mathbf{A}^T)$. \square

To summarise, the last four corollaries provide a standard way of finding orthonormal bases for the four fundamental subspaces.

THEOREM 5.2. Defining the SVD of \mathbf{A} as in Eq. (5.9) as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = [\mathbf{U}_1 \ \mathbf{U}_2] \mathbf{\Sigma} [\mathbf{V}_1 \ \mathbf{V}_2]^T$$

It follows that

- \mathbf{U}_1 forms a basis for the column space $\mathcal{R}(\mathbf{A})$
- \mathbf{U}_2 forms a basis for the left null space $\mathcal{N}(\mathbf{A}^T)$
- \mathbf{V}_1 forms a basis for the row space $\mathcal{R}(\mathbf{A}^T)$
- \mathbf{V}_2 forms a basis for the null space $\mathcal{N}(\mathbf{A})$

5.3 Application 2: Low rank approximation

We have seen that SVD can be written as a sum of rank one matrices

$$\mathbf{A} = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

If the rank of \mathbf{A} is r , then $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$ therefore

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

Sometimes we want to store smaller $\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}$ matrices e.g. in the computer memory so we want to approximate \mathbf{A} using the first (largest) k singular values

$$\mathbf{A}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (5.10)$$

☞ **Comment:** The low order singular values $(\sigma_1, \sigma_2, \dots)$ carry more information than the high order ones $(\sigma_r, \sigma_{r-1}, \dots)$.

The figure below shows how one approximation of the sum above is obtained as a product of a (column) \times (singular value) \times (row)

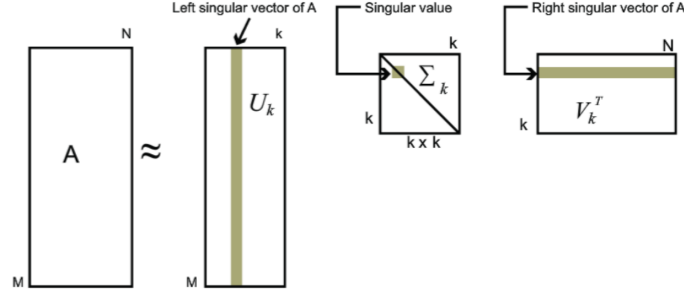


Figure 12: One term of the approximation sum in Eq. (5.10).

5.3.1 SVD and norms

To start with, an interesting result is that SVD provides the best rank k approximation of a matrix \mathbf{A} . The following is the 2-norm case of the Eckart-Young-Mirsky theorem. We will prove it for the 2-norm first and generalise for the Frobenius norm later [23]. We denote $\|\cdot\|$ or $\|\cdot\|_2$ the 2-norm.

THEOREM 5.3. *Let \mathbf{A} be an $m \times n$ matrix of rank r , $k < r$ and \mathbf{A}_k the rank k SVD of \mathbf{A} . Then for any matrix \mathbf{X} with $\text{rank}(\mathbf{X}) = k$*

$$\min_{\text{rank}(\mathbf{X})=k} \|\mathbf{A} - \mathbf{X}\| = \sigma_{k+1} \quad (5.11)$$

The equality is attained by

$$\mathbf{X} = \mathbf{A}_k = \sum_{i=1}^k \mathbf{u}_i \sigma_i \mathbf{v}_i^T \quad (5.12)$$

Proof.

[24], [25], [26] Since $\text{rank}(\mathbf{X}) = \dim(\mathcal{R}(\mathbf{A})) = k$, we have $\dim(N(\mathbf{A})) = n - k$. Also, $\dim(\mathbf{V}_{k+1}) = k + 1$, $\mathbf{V}_k = [\mathbf{v}_1 \ \dots \ \mathbf{v}_{k+1}]$ are the first $k + 1$ right singular vectors of \mathbf{A} , it must be true one (call it \mathbf{z}) falls in $N(\mathbf{X})$, i.e.

$$\exists \mathbf{z} \in \mathbb{R}^n : \mathbf{z} \in N(\mathbf{X}) \cap \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\}, \quad \|\mathbf{z}\|_2 = 1$$

$\|\mathbf{z}\|_2 = 1$ because of course \mathbf{V} is an orthogonal matrix. We can write

$$\begin{aligned} \mathbf{z} &= c_1 \mathbf{v}_1 + \dots + c_{k+1} \mathbf{v}_{k+1} \\ \therefore \|\mathbf{z}\| &= 1 \Rightarrow \sum_{i=1}^{k+1} c_i^2 = 1 \end{aligned} \quad (1)$$

Using one of the properties of the 2-norm and because $\mathbf{z} \in N(\mathbf{X})$:

$$\|\mathbf{A} - \mathbf{X}\| \geq \|(\mathbf{A} - \mathbf{X})\mathbf{z}\| = \|\mathbf{A}\mathbf{z}\| \quad (2)$$

If we expand $\mathbf{A}\mathbf{z}$ using the rank- k SVD of \mathbf{A} and Eq. (1) we have:

$$\mathbf{A}\mathbf{z} = \left(\sum_{i=1}^{k+1} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \right) \left(\sum_{j=1}^{k+1} c_j \mathbf{v}_j \right) = \sum_{i=1}^{k+1} \sigma_i c_i \mathbf{u}_i \mathbf{v}_i^T \mathbf{v}_i = \sum_{i=1}^{k+1} \sigma_i c_i \mathbf{u}_i \quad (3)$$

Using the orthonormality of \mathbf{u}_i and the ordering of σ_i values:

$$\|\mathbf{A}\mathbf{z}\|^2 = \sum_{i=1}^{k+1} \sigma_i^2 c_i^2 \geq \sigma_{k+1}^2 \underbrace{\sum_{i=1}^{k+1} c_i^2}_{=\|\mathbf{z}\|=1} = \sigma_{k+1}^2 \quad (4)$$

Therefore Eq. (2) with the aid of Eq. (3) and Eq. (4) yields the following simple expression

$$\|\mathbf{A} - \mathbf{X}\| = \max_{\|\mathbf{v}\|=1} \|(\mathbf{A} - \mathbf{X})\mathbf{v}\| \geq \|(\mathbf{A} - \mathbf{X})\mathbf{z}\| = \|\mathbf{A}\mathbf{z}\| \geq \sigma_{k+1}^2$$

Now we want to show that the equality for minimum is obtained for $\mathbf{X} = \mathbf{A}_k = \sum_{i=1}^k \mathbf{u}_i \sigma_i \mathbf{v}_i^T$.

$$\mathbf{A} - \mathbf{A}_k = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T - \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=k+1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

This is the SVD of $\mathbf{A} - \mathbf{A}_k$ with largest singular value σ_{k+1} so by definition we have $\|\mathbf{A} - \mathbf{A}_k\|_2 = \max \left(\sum_{i=k+1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T \right) = \sigma_{k+1}$. To conclude, \mathbf{A}_k has the best norm-2 approximation. \square

TODO: Eckart-Young-Mirsky theorem generalisation to be written in the future.

5.4 Application 3: Image compression



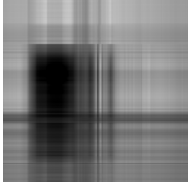


As mentioned before, we can only pick the first k largest singular values along with the first k left and right singular vectors to compress an matrix, e.g. image using less values. Particularly, if the image is $m \times n$, then we only need $m \times k + k + k \times n$ numbers to approximate the image using the SVD product. To actually save data, we want [27]

$$k(m + n + 1) < mn \Rightarrow k < \frac{mn}{m + n + 1}$$

Therefore SVD doesn't always save data, which is one of its disadvantages compares to other image compression methods such as JPG.

A Matlab script that performs SVD image compression is listed in App. A.9. Some of the results it generates are shown in the table below. Note that the rank of the cameraman image is 253 and its size is 256×256 . Also, notice how Eckart-Young-Mirsky theorem is confirmed.

Table 1: Approximations of the cameraman image using SVD for various ranks.

Image	Rank	Output	Numbers re- quired	Norm 2 of difference	Singular value r+1
Original	253		65536		
Approximated	1		513	28.383	28.383
Approximated	2		1026	21.506	21.506
Approximated	32		16416	2.379	2.379
Approximated	64		32832	1.157	1.157

5.4.1 Application 3.1: Image deblurring

TODO: maybe in future

http://www.mathcs.emory.edu/~nagy/courses/fall06/ID_lecture1.pdf

5.5 Application 4: Pseudoinverse and least squares

In App. A.10, we studied the system $\mathbf{A}_{m \times n} \mathbf{x} = \mathbf{b}$, $m > n$. $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$. Therefore in this case we have to "fit" $\mathbf{b} \in \mathbb{R}^m$ in the smaller space \mathbb{R}^n , spanned by the columns of \mathbf{A} . We "fit" \mathbf{b} by finding its smallest distance point in the column space of \mathbf{A} , which is its projection on it. So instead of solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ we solve $\mathbf{A}\bar{\mathbf{x}} = \mathbf{p}$, where \mathbf{p} is the projection of \mathbf{b} in the column space of \mathbf{A} .

To reiterate, we found that:

1. $\bar{\mathbf{x}}$ is a least squares solution of $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$ iff $\bar{\mathbf{x}}$ is a solution of the normal equations

$$(\mathbf{A}^T \mathbf{A}) \bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b} \quad (5.13)$$

2. \mathbf{A} has linearly independent columns if and only if $\mathbf{A}^T \mathbf{A}$ is invertible. In this case, the least squares solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ is unique and is given by

$$\bar{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (5.14)$$

☞ **Comment:** If the stability of the system is in doubt therefore we cannot invert $\mathbf{A}^T \mathbf{A}$, then we can solve the system by QR factorisation or SVD. We will study the latter. QR factorisation is essentially a compact matrix form of the Gram Schmidt decomposition, where \mathbf{Q} is orthogonal and \mathbf{R} is upper triangular. The QR solution is proven in App. A.11

Matrix $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ is called pseudoinverse of \mathbf{A} and we will show that it can be readily computed by SVD. The classical inverse exists only if \mathbf{A} is $n \times n$ and $\text{rank}(\mathbf{A}) = n$. The pseudoinverse \mathbf{A}^+ exists for all $\mathbf{A}_{m \times n}$ matrices. It can be thought as a generalisation of the inverse. When $\mathbf{A}\mathbf{x} = \mathbf{b}$ has no solutions, which is the case we study, it provides the matrix such that $\mathbf{A}^+ \mathbf{x}$ is closest to \mathbf{b} . If many solutions exist, it picks the solution that minimises $\|\mathbf{x}\|_2$.

Pseudoinverse does not necessarily have the strong $\mathbf{A}^+ \mathbf{A} = \mathbf{I}$ property, but satisfies all of the following properties [28].

COROLLARY 5.7.

1. $\mathbf{A}\mathbf{A}^+ \mathbf{A} = \mathbf{A}$ ($\mathbf{A}\mathbf{A}^+$ does not have to be the identity matrix but maps all column vectors of \mathbf{A} to themselves)
2. $\mathbf{A}^+ \mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$
3. $(\mathbf{A}\mathbf{A}^+)^T = \mathbf{A}\mathbf{A}^+$
4. $(\mathbf{A}^+ \mathbf{A})^T = \mathbf{A}^+ \mathbf{A}$
5. $\text{rank}(\mathbf{A}^+) = \text{rank}(\mathbf{A})$

Also, note that $\mathbf{A}\mathbf{A}^+$ and $\mathbf{A}^+ \mathbf{A}$ are as close to \mathbf{I}_m and \mathbf{I}_n respectively as possible.

It can easily be proven [29] that the solution $\bar{\mathbf{x}}$ in Eq. (5.14) in terms of singular vectors and values is given by:

$$\bar{\mathbf{x}} = \mathbf{V}\mathbf{D}^+ \mathbf{U}^T \mathbf{b}, \quad \mathbf{D}^+ = (\mathbf{D}^2)^{-1} \mathbf{D} \quad (5.15)$$

In Eq. (5.15), matrix \mathbf{A} doesn't have to be invertible so that's a general way to compute its pseudo inverse. \mathbf{D}^+ is of course the pseudo inverse of \mathbf{D} . Consider the following cases when finding $\bar{\mathbf{x}}$ [30].

- $\text{rank}(\mathbf{A}) = n$, i.e. $\sigma_1 \geq \dots \geq \sigma_n > 0$. Then \mathbf{D} 's pseudoinverse coincides with its inverse so it's trivial to find:

$$\mathbf{D}^+ = \mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & & \\ & \ddots & \\ & & \frac{1}{\sigma_n} \end{bmatrix}$$

Then the matrix $\mathbf{V}\mathbf{D}^+ \mathbf{U}^T$ from Eq. (5.15) is indeed the pseudoinverse of \mathbf{A} as

$$\mathbf{A}^T \mathbf{A} = \mathbf{V}\mathbf{D}^+ \mathbf{U}^T \mathbf{U} \mathbf{D} \mathbf{V}^T = \mathbf{I}_n$$

- $\text{rank}(\mathbf{A}) = r < n$, therefore \mathbf{A}^{-1} does not exist. However, the pseudoinverse does exist and for each column we have:

$$\mathbf{A}^+ \mathbf{u}_i = \begin{cases} \frac{1}{\sigma_i} \mathbf{v}_i, & i \leq r = \text{rank}(\mathbf{A}) \\ 0, & \text{otherwise} \end{cases}$$

In matrix form, this is written as

$$\mathbf{A}^+ = \mathbf{V} \mathbf{D}^+ \mathbf{U}^T, \quad \mathbf{D}^+ = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \quad (5.16)$$

Again, $\mathbf{D}^+ \mathbf{D}$ is as close to the identity as possible.

5.6 Application 5: Principal Component Analysis (PCA)

PCA is a statistical procedure concerned with elucidating the covariance structure of a set of variables. In particular it allows us to identify the principal directions in which the data varies. To be more precise, [31],

- it projects data onto a lower-dimensional manifold that preserves as much information as possible,
- minimises reconstruction error,
- maximises the variance (signal) of the projected data,
- maximises the mutual information between original and projected data

TODO!

<http://www.cse.psu.edu/~rtc12/CSE586Spring2010/lectures/pcaLectureShort.pdf>

<https://www.ini.rub.de/PEOPLE/wiskott/Teaching/Material/PrincipalComponentAnalysis-LectureNotesPublic.pdf>

<https://alliance.seas.upenn.edu/~cis520/wiki/index.php?n=Lectures.PCA>

<https://courses.cs.washington.edu/courses/cse521/16sp/521-lecture-9.pdf>

http://pillowlab.princeton.edu/teaching/statneuro2018/slides/notes04_PCA.pdf

https://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf

http://labs.seas.wustl.edu/bme/raman/Lectures/Lecture11_PCA.pdf

<https://www.cs.cmu.edu/~elaw/papers/pca.pdf>

<http://www.cis.upenn.edu/~cis515/cis515-12-sl13.pdf>

<http://luthuli.cs.uiuc.edu/~daf/courses/CS-498-DAF-PS/Lecture%209%20-%20PCA.pdf>

http://www.sci.utah.edu/~shireen/pdfs/tutorials/Elhabian_PCA09.pdf

verbose explanation here:

<https://stats.stackexchange.com/questions/2691/making-sense-of-principal-component-analysis-eigenvectors>

quick summary:

<https://stats.stackexchange.com/questions/305430/why-do-we-want-to-maximize-the-variance-in-principal-component-analysis>

<https://medium.com/@aptrishu/understanding-principle-component-analysis-e32be0253ef0>

5.6.1 Application of PCA for facial recognition

TODO: this is an optional section, not sure if I will write it https://www.cc.gatech.edu/~isbell/reading/papers/draper_cvui03.pdf <https://digitalrepository.trincoll.edu/cgi/viewcontent.cgi?article=1221&context=theses>

A Appendices



A.1 Appendix 1. Matrix form of general conic section equation

TODO: not urgent, do it in the future <http://calvino.polito.it/~salamon/P/L/lag21.pdf>
<https://zdaugherty.ccny.cuny.edu/teaching/m202s16/resources/conic-rotation.pdf>
[https://math.stackexchange.com/questions/1447730/drawing-ellipse-from-eigenvalue-eigenvector?
utm_medium=organic&utm_source=google_rich_qa&utm_campaign=google_rich_qa](https://math.stackexchange.com/questions/1447730/drawing-ellipse-from-eigenvalue-eigenvector?utm_medium=organic&utm_source=google_rich_qa&utm_campaign=google_rich_qa)

A.2 Appendix 2. 3D paraboloid equations

<http://tutorial.math.lamar.edu/Classes/CalcIII/QuadricSurfaces.aspx>

TODO: not urgent, do it in the future

A.3 Appendix 3. Cholesky factorisation

Cholesky factorisation (or decomposition) is a particular case of LU decomposition if our matrix satisfies some conditions.

LEMMA A.1. *Every symmetric positive definite matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ has a unique factorisation $\mathbf{A} = \mathbf{L}^T \mathbf{L}$, where \mathbf{L} is a lower triangular matrix whose diagonal is positive ($l_{ii} > 0$).*

Proof. [32]

- Fact 1. If \mathbf{L} is a unit lower triangular matrix and \mathbf{U} is upper triangular, such that $\mathbf{A} = \mathbf{LU}$, then \mathbf{L} and \mathbf{U} are unique, i.e. if $\mathbf{A} = \tilde{\mathbf{L}} \tilde{\mathbf{U}} \tilde{\mathbf{L}}$ is a unit lower triangular matrix and $\tilde{\mathbf{U}}$ is upper triangular, then $\mathbf{L} = \tilde{\mathbf{L}}$, $\mathbf{U} = \tilde{\mathbf{U}}$.
- Fact 2. If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric positive definite and $\mathbf{A} = \mathbf{LU}$ then the diagonal entries of \mathbf{U} are positive. *Proof sketch:* Use that $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \forall \mathbf{x} \in \mathbb{R}^n$. Define $\mathbf{e}_1 \in \mathbb{R}^n : \mathbf{e}_1 = [1 \ 0 \ \dots \ 0]^T$, ..., $\mathbf{e}_i \in \mathbb{R}^n : \mathbf{e}_i = [0 \ \dots \ 1 \ \dots \ 0]^T$, the i -th column of the unit matrix. Then $\mathbf{e}_i^T \mathbf{A} \mathbf{e}_i = 1 \cdot a_{ii} \cdot 1$ etc.
- Fact 3. An upper triangular matrix \mathbf{U} can be written as the product of a diagonal matrix \mathbf{D} with a unit upper triangular (unit on the diagonal and upper triangular at the same time) $\tilde{\mathbf{U}}$, i.e. $\mathbf{U} = \mathbf{D} \tilde{\mathbf{U}}$ with:

$$\mathbf{D} = \begin{bmatrix} u_{11} & & & & \\ & u_{22} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & u_{nn} \end{bmatrix} \quad (A.1)$$

$$\tilde{\mathbf{U}} = \begin{bmatrix} 1 & u_{12}/u_{11} & & \dots & \dots & u_{1n}/u_{11} \\ & 1 & u_{23}/u_{22} & \dots & \dots & u_{2n}/u_{22} \\ & & 1 & u_{34}/u_{33} & \dots & u_{3n}/u_{33} \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & 1 \end{bmatrix}$$

Therefore,

$$\mathbf{A} = \mathbf{LU} = \mathbf{LD} \tilde{\mathbf{U}}$$

Also,

$$\mathbf{A}^T = \mathbf{LU} = (\mathbf{LD} \tilde{\mathbf{U}})^T = \tilde{\mathbf{U}}^T \mathbf{D} \mathbf{L}^T$$

We know that \mathbf{A} is symmetric and (Fact 1) that the LU decomposition is unique, therefore:

$$\mathbf{LU} = \tilde{\mathbf{U}}^T \mathbf{D} \mathbf{L}^T = \tilde{\mathbf{U}}^T (\mathbf{D} \mathbf{L}^T) \quad (\text{lower unit triangular} \times \text{upper triangular}) \implies \quad (A.2)$$

$$\mathbf{L} = \tilde{\mathbf{U}}^T, \mathbf{U} = \mathbf{D} \mathbf{L}^T$$

From Eq. (A.1), Eq. (A.2) we have found that the LU decomposition of the symmetric, positive definite \mathbf{A} can be written as:

$$\mathbf{A} = \mathbf{LDL}^T$$

From Fact 2 we know that the entries of \mathbf{D} are positive, so we can define

$$\mathbf{D}^{1/2} = \begin{bmatrix} \sqrt{u_{11}} & & & & \\ & \sqrt{u_{22}} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \sqrt{u_{nn}} \end{bmatrix}$$

If we define $\mathbf{R} = \mathbf{D}^{1/2} \mathbf{L}^T$, then \mathbf{A} can be written as

$$\begin{aligned}
\mathbf{A} &= \mathbf{L}\mathbf{D}\mathbf{L}^T = \mathbf{L}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{L}^T \\
&= (\mathbf{D}^{1/2}\mathbf{L}^T)^T \mathbf{D}^{1/2}\mathbf{L}^T \\
&= \mathbf{R}^T \mathbf{R}
\end{aligned}$$

That is the *Cholesky factorisation*.

□

EXAMPLE A.1. Find the Cholesky factorisation of $\mathbf{A} = \begin{bmatrix} 4 & -4 & 8 \\ -4 & 8 & -4 \\ 8 & -4 & 29 \end{bmatrix}$ [33].

Solution:

$$\begin{aligned}
\begin{bmatrix} 4 & -4 & 8 \\ -4 & 8 & -4 \\ 8 & -4 & 29 \end{bmatrix} &= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} 4 & -4 & 8 \\ 0 & 4 & 4 \\ 0 & 0 & 9 \end{bmatrix}}_{\mathbf{U}} \\
&= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{\tilde{\mathbf{U}}=\mathbf{L}^T} \\
&= \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{\mathbf{D}^{1/2}} \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{\mathbf{D}^{1/2}} \underbrace{\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{L}^T} \\
&= \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ -2 & 2 & 0 \\ 4 & 2 & 3 \end{bmatrix}}_{\mathbf{R}^T} \underbrace{\begin{bmatrix} 2 & -2 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}}_{\mathbf{R}}
\end{aligned}$$

■

A.4 Properties of the gradient operator

COROLLARY A.1. (Property 1) For the vector field of the gradient, its product with any unit vector \mathbf{v} is equal to the derivative along \mathbf{v} , i.e. the projection of the gradient on any unit vector \mathbf{v} gives the directional derivative along that direction

$$(\nabla f(x)) \cdot \mathbf{v} = D_{\mathbf{v}} f(x) \quad (\text{A.3})$$

EXAMPLE A.2. Consider the function $f(x, y) = x \cos y$. Find its directional derivative along $\mathbf{v} = (0, 1)$ and along $\mathbf{v} = (2, 1)$ [34].

Solution:

Find the gradient in terms of cartesian coordinates first

$$\nabla f = \cos y \mathbf{i} - x \sin y \mathbf{j}$$

$\mathbf{v} = (0, 1)$ is already a unit vector so its directional derivative along it is

$$(\nabla f(x, y)) \cdot \mathbf{v} = -x \sin y$$

$\mathbf{v} = (2, 1)$ becomes $\hat{\mathbf{v}} = \frac{2}{\sqrt{5}} \mathbf{i} + \frac{1}{\sqrt{5}} \mathbf{j}$ so the directional derivative is

$$(\nabla f(x, y)) \cdot \hat{\mathbf{v}} = (\cos y, -x \sin y) \cdot \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right) = \frac{1}{\sqrt{5}} (2 \cos y - x \sin y)$$

COROLLARY A.2. (Property 2) Suppose f is a differentiable function of two or three variables. The maximum value of the directional derivative $D_{\mathbf{u}}$ is $|\nabla f(\mathbf{x})|$ and it occurs when \mathbf{u} has the same direction as the gradient vector $\nabla f(\mathbf{x})$.

Proof. From Eq. (A.3) we have:

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u} = |\nabla f| \underbrace{|\mathbf{u}|}_{1} \cos(\nabla f, \mathbf{u}) \leq |\nabla f|$$

The equality occurs when $\cos(\nabla f, \mathbf{u}) = 1$, i.e. when the angle between ∇f and \mathbf{u} is 0. □

COROLLARY A.3. (Property 3) Let $w = f(x, y, z)$ be a differentiable function of three variables. At any point $P(x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$, the gradient $\nabla f|_P$ is perpendicular to the surface. By this we mean it's perpendicular to the tangent to any curve that lies on the surface and goes through P .

Proof. [35] Parametrise $x := x(t), y := y(t), z := z(t)$ and let

$$\mathbf{r}(t) = (x(t), y(t), z(t))$$

be a curve on the level surface with $\mathbf{r}(t_0) = (x(t_0), y(t_0), z(t_0))$. Let $g(t) := f(x(t), y(t), z(t))$. Since the curve is on the level surface we have $g(t) = f(x(t), y(t), z(t)) = c$ and differentiating this w.r.t t gives:

$$\frac{dg}{dt} = \frac{dg}{dx} \bigg|_P \frac{dx}{dt} \bigg|_{t_0} + \frac{dg}{dy} \bigg|_P \frac{dy}{dt} \bigg|_{t_0} + \frac{dg}{dz} \bigg|_P \frac{dz}{dt} \bigg|_{t_0} = 0$$

In vector form:

$$\left(\frac{dg}{dx} \bigg|_P, \frac{dg}{dy} \bigg|_P, \frac{dg}{dz} \bigg|_P \right) \cdot \left(\frac{dx}{dt} \bigg|_{t_0}, \frac{dy}{dt} \bigg|_{t_0}, \frac{dz}{dt} \bigg|_{t_0} \right) \Rightarrow \nabla f|_P \cdot \mathbf{r}'(t_0) = 0$$

□

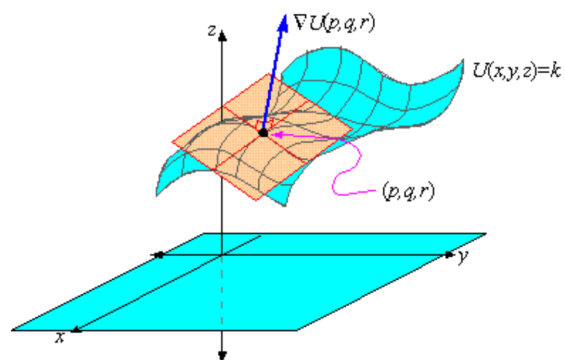


Figure 13: Gradient is perpendicular to the local tangential plane.

A.5 Taylor series for multi variable functions

DEFINITION A.1. An n -times differentiable real function $f : \mathbb{R} \rightarrow \mathbb{R}$ can be expanded in the form of Taylor series about a point $x = x_0$ as

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \frac{f^{(3)}(x_0)}{3!}(x - x_0)^3 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \epsilon_n \quad (\text{A.4})$$

, where ϵ_n is the remainder or error polynomial.

Therefore we can approximate the function f using first order terms at x_0 as $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$, using second order terms as $f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2$, etc. Alternatively, when Δx is small, e.g. the 2nd degree approximation at $x = x_0 + \Delta x$ can be written as:

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x + \frac{f''(x_0)}{2!}\Delta x^2$$

For a function of two independent variables $f(x, y)$, if we want to expand about (x_0, y_0) using 2nd derivatives we take partial derivatives letting $f_x = \frac{\partial f}{\partial x}$, $f_y = \frac{\partial f}{\partial y}$, $x = x_0 + \Delta x$, $y = y_0 + \Delta y$ in the above form and have [36]

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) \approx & f(x_0, y_0) + f_x \Delta x + f_y \Delta y + \\ & f_{xx}(x_0, y_0) \frac{\Delta x^2}{2} + f_{yy}(x_0, y_0) \frac{\Delta y^2}{2} + f_{xy}(x_0, y_0) \Delta x \Delta y \end{aligned}$$

Equivalently, using the first form [37]:

$$\begin{aligned} f(x, y) \approx & f(x_0, y_0) + f_x \cdot (x - x_0) + f_y \cdot (y - y_0) + \\ & f_{xx}(x_0, y_0) \frac{(x - x_0)^2}{2} + f_{yy}(x_0, y_0) \frac{(y - y_0)^2}{2} + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) \end{aligned}$$

A bit more elegantly, denoting the small displacements as $\Delta x := x - x_0$, $\Delta y := y - y_0$:

$$\begin{aligned} f(x, y) \approx & f(x_0, y_0) + f_x \cdot \Delta x + f_y \cdot \Delta y + \\ & \frac{1}{2} (f_{xx}(x_0, y_0)\Delta x^2 + f_{yy}(x_0, y_0)\Delta y^2 + 2f_{xy}(x_0, y_0)\Delta x \Delta y) \end{aligned} \quad (\text{A.5})$$

We want to find a more compact matrix-based way to write the above equations denoting our 2-variable function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ as $f(\mathbf{x})$, $\mathbf{x} = [x \ y]^T$ and seeking the 2nd order approximation around a fixed point $\mathbf{x}_0 = [x_0 \ y_0]^T$. From Section 1 we know that the sum $f_{xx}(x_0, y_0) \frac{(x - x_0)^2}{2} + f_{yy}(x_0, y_0) \frac{(y - y_0)^2}{2} + f_{xy}(x_0, y_0)(x - x_0)(y - y_0)$ can be written as a quadratic form, particularly:

$$\begin{aligned} f_{xx}(x_0, y_0) \frac{(x - x_0)^2}{2} + f_{yy}(x_0, y_0) \frac{(y - y_0)^2}{2} + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) = \\ \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \begin{bmatrix} f_{xx}(\mathbf{x}_0) & f_{xy}(\mathbf{x}_0) \\ f_{yx}(\mathbf{x}_0) & f_{yy}(\mathbf{x}_0) \end{bmatrix} (\mathbf{x} - \mathbf{x}_0) = \\ \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) = \end{aligned}$$

, where $\mathbf{H}_f(\mathbf{x}_0)$ is the Hessian matrix of f at fixed point \mathbf{x}_0 .

Recall that the gradient of f is $\nabla f = [f_x \ f_y]$ so the term can be written as the dot product $(\mathbf{x} - \mathbf{x}_0)^T \cdot \nabla f(\mathbf{x}_0)$.

COROLLARY A.4. Putting everything together, the 2nd order Taylor approximation of a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ or $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in general around a fixed point \mathbf{x}_0 is

$$f(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \cdot \nabla f(\mathbf{x}_0) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) \quad (\text{A.6})$$

, where $\mathbf{H}_f(\mathbf{x}_0)$ is the Hessian matrix ($n \times n$) evaluated at \mathbf{x}_0 . The Hessian matrix is of course symmetric for any size as $f_{xy} = f_{yx}$.

The following result is now straightforward from Eq. (A.6).

COROLLARY A.5. Let $f(x, y)$ be a twice differentiable function, and let (x_0, y_0) be a critical point for f , i.e. $\nabla f(x_0, y_0) = \mathbf{0}$.

1. If $\mathbf{H}_f(x_0, y_0)$ is positive definite, then x_0, y_0 is a local minimum for f .
2. If $\mathbf{H}_f(x_0, y_0)$ is negative definite, then x_0, y_0 is a local maximum for f .
3. If $\mathbf{H}_f(x_0, y_0)$ is indefinite, then x_0, y_0 is a saddle point for f .

Proof. [38] We know that at the critical point the gradient vanishes, i.e.

$$\nabla f(x_0, y_0) = \mathbf{0}$$

So the 2nd order Taylor expansion at the critical point $\mathbf{x}_0 = (x_0, y_0)$ looks like

$$f(\mathbf{x}_0 + \Delta \mathbf{x}) \approx f(\mathbf{x}_0) + \frac{1}{2} \Delta \mathbf{x}^T \mathbf{H}_f(\mathbf{x}_0) \Delta \mathbf{x} \quad (1)$$

Thus, for small displacements Δx , the Hessian tells us how the function behaves around the critical point.

1. If the Hessian is positive definite, then $\Delta \mathbf{x}^T \mathbf{H}_f(\mathbf{x}_0) \Delta \mathbf{x} > 0$ therefore from Eq. (1) we have $f(\mathbf{x}_0 + \Delta \mathbf{x}) > f(\mathbf{x}_0)$ around \mathbf{x}_0 . Then in every direction around that point $f(\mathbf{x})$ grows thus \mathbf{x}_0 is a local minimum.
2. If the Hessian is negative definite, then $\Delta \mathbf{x}^T \mathbf{H}_f(\mathbf{x}_0) \Delta \mathbf{x} < 0$ therefore from Eq. (1) we have $f(\mathbf{x}_0 + \Delta \mathbf{x}) < f(\mathbf{x}_0)$ around \mathbf{x}_0 , i.e. \mathbf{x}_0 . Then in every direction around that point $f(\mathbf{x})$ shrinks thus \mathbf{x}_0 is a local maximum.
3. If the Hessian is indefinite, then there exists $\Delta \mathbf{x} \neq \mathbf{0}$ such that $\Delta \mathbf{x}^T \mathbf{H}_f(\mathbf{x}_0) \Delta \mathbf{x} = 0$. In this case the test fails; along this direction we aren't really sure whether the function f is increasing or decreasing as we move away from \mathbf{x}_0 ; our second order approximation isn't good enough and we need higher order data to decide.

When considering the definitiveness of the Hessian, remember the equivalent definitiveness criteria from Theorems 3.1, 3.3. We will use those to formulate an equivalent 2nd partial derivative extremum test. \square

COROLLARY A.6. Let f be twice differentiable with continuous second derivatives around (x_0, y_0) and (x_0, y_0) be a critical point, i.e. $\nabla f(x_0, y_0) = \mathbf{0}$. The determinant of the Hessian is $\det(\mathbf{H}_f)(x_0, y_0) = f_{xx}f_{yy} - f_{xy}f_{yx} = f_{xx}f_{yy} - f_{xy}^2$.

1. If $\det(\mathbf{H}_f)(x_0, y_0) > 0$ and $f_{xx} > 0$ then x_0, y_0 is a local minimum.
2. If $\det(\mathbf{H}_f)(x_0, y_0) > 0$ and $f_{xx} < 0$ then x_0, y_0 is a local maximum.
3. If $\det(\mathbf{H}_f)(x_0, y_0) < 0$ then x_0, y_0 is a saddle point.
4. If $\det(\mathbf{H}_f)(x_0, y_0) = 0$ then the test is inconclusive and (x_0, y_0) can be either a minimum, maximum, or saddle point.

Proof.

1. From Theorem 3.1, this proposition implies that the determinants of all upper left principal minors are positive (as $\det(\mathbf{H}_f)(x_0, y_0) > 0$ and $f_{xx} > 0$ (which implies also $f_{yy} > 0$) therefore $\mathbf{H}_f(x_0, y_0)$ is positive definite, thus x_0, y_0 is a local minimum.
2. From Theorem 3.3 this proposition implies that the determinants of all upper left principal minors have sign $(-1)^k$ as $\det(\mathbf{H}_f)(x_0, y_0) > 0$ and $f_{xx} < 0$ (which implies also $f_{yy} > 0$) therefore $\mathbf{H}_f(x_0, y_0)$ is negative definite, thus x_0, y_0 is a local maximum.
3. If $\det(\mathbf{H}_f)(x_0, y_0) < 0$ then we have one positive and one negative eigenvalue therefore $\mathbf{H}_f(x_0, y_0)$ is indefinite, thus x_0, y_0 is a saddle point.

An intuitive way to remember whether the sign of f_{xx} combined with positive determinant implies minimum or maximum is to compare it with the 1D local minimum case. \square

Proof.

[39] This is an alternative proof that does not consider the definitiveness criteria. From Eq. (A.5) with $\Delta x = x - x_0$, $\Delta y = y - y_0$ around the critical point (x_0, y_0) , i.e. $\nabla f(x_0, y_0) = \mathbf{0}$, we have

$$\begin{aligned} \Delta f &= f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \\ &= \frac{1}{2} [f_{xx}(\Delta x)^2 + f_{yy}(\Delta y)^2 + 2f_{xy}\Delta x\Delta y] \end{aligned} \quad (2)$$

Maximum: What condition do we need for a local maximum?

For a maximum, all small steps away from (x_0, y_0) will lead to $\Delta f < 0$ therefore Eq. (2) implies that

$$f_{xx}(\Delta x)^2 + f_{yy}(\Delta y)^2 + 2f_{xy}\Delta x\Delta y < 0 \quad (3)$$

for arbitrary sufficiently small $\Delta x, \Delta y$. We can therefore set in Eq. (3) $\Delta x = 0$, which yields:

$$f_{xx} < 0 \quad (4)$$

(Also similarly, $f_{yy} < 0$.) We can also align the displacements as $\Delta x = \lambda\Delta y$, i.e. in Eq. (3) set $(\Delta x, \Delta y) := (\lambda\Delta y, \Delta y)$, which yields:

$$\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} < 0 \quad \forall \lambda \neq 0$$

Multiplying by f_{xx} (remember Eq. (4)) we obtain:

$$\begin{aligned} \lambda^2 f_{xx}^2 + 2\lambda f_{xy}f_{xx} + f_{yy}f_{xx} &> 0 \Rightarrow \\ f_{xy}^2 - f_{xx}f_{yy} &< (f_{xx}\lambda + f_{xy})^2 \quad \forall \lambda \neq 0 \end{aligned} \quad (5)$$

Picking $\lambda = -f_{xy}/f_{xx}$ we can set the right side to zero so we end up with:

$$f_{xy}^2 - f_{xx}f_{yy} < 0 \quad (6)$$

The conditions in Eq. (4) and Eq. (6) state that (x_0, y_0) is a local maximum when:

$$\begin{aligned} f_{xx} < 0 \quad \text{and} \quad & f_{xy}^2 - f_{xx}f_{yy} < 0 \Leftrightarrow \\ & f_{xx}f_{yy} - f_{xy}^2 > 0 \Leftrightarrow \\ & \det(\mathbf{H}_f)(x_0, y_0) > 0 \end{aligned} \quad (7)$$

Using the Principal Minor equivalence criterion (Theorem 3.3), Eq. (7) implies that \mathbf{H}_f must be negative definite at at local maximum (x_0, y_0) .

Minimum: The same type of argument can be applied to formulate the condition for a minimum.

Saddle point:

For a saddle point, we'd have $\Delta f = 0$ in Eq. (2). However, we don't know how Δx and Δy change as we move away from (x_0, y_0) . We can still set $\Delta x = \lambda\Delta y$ but λ needs to be found. Again, from Eq. (3) we have:

$$\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} = 0 \quad (8)$$

This implies that there are only two direction λ_1, λ_2 for which $\Delta x = \lambda_i\Delta y$ obtained by solving the above equation. Multiplying by f_{xx} and completing the square:

$$\begin{aligned} \lambda^2 f_{xx}^2 + 2\lambda f_{xy}f_{xx} + f_{yy}f_{xx} &= f_{yy}^2 - f_{xx}f_{yy} \Rightarrow \\ (\lambda f_{xx} + f_{yy})^2 &= f_{xy}^2 - f_{xx}f_{yy} \end{aligned}$$

Because λ solves Eq. (8), it cannot satisfy $(\lambda f_{xx} + f_{yy})^2 = 0$ at the same time therefore

$$(\lambda f_{xx} + f_{yy})^2 > 0 \Rightarrow f_{xy}^2 - f_{xx}f_{yy} > 0 \quad (9)$$

To sum up, the condition for saddle point at (x_0, y_0) is given by Eq. (9). \square

A.6 Intuition behind Lagrange multiplier

TODO: maybe in the future

https://www.reddit.com/r/math/comments/g4w7m/can_anyone_provide_a_concise_and_intuitive/

<https://ocw.mit.edu/courses/mathematics/18-02sc-multivariable-calculus-fall-2010/2.-partial-derivatives>

[part-c-lagrange-multipliers-and-constrained-differentials/session-40-proof-of-lagrange-multipliers/](#)

[MIT18_02SC_notes_22.pdf](#)

<http://www.uoguelph.ca/economics/repec/workingpapers/2010/2010-03.pdf>

A.7 Relationship between rank and number of non-zero eigenvalues

The relationship is derived in a number of sets and a special case holds for symmetric matrices.

- Eigenvectors of distinct eigenvalues are lin. ind.
- Eigenvectors of nonzero eigenvalues of \mathbf{A} are in its column space (or range) $\mathcal{R}(\mathbf{A})$ as $\mathbf{v}_i = \mathbf{A}\mathbf{v}_i/\lambda_i$ is a combination of \mathbf{A} 's columns.
- Therefore, if $r := \text{rank}(\mathbf{A})$ and we have k (not necessarily distinct) eigenvalues, then $k \leq r$. Also for the null space, $\dim(\mathcal{N}(\mathbf{A})) = n - r \geq n - k$.
- If \mathbf{A} is *symmetric*, all its eigenvalues are distinct therefore all its eigenvectors are lin. ind. so their number equals the dimension of their spanning set, i.e. $k = \dim(\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\})$. But each eigenvector is in the column space of \mathbf{A} so it follows $k = r = \text{rank}(\mathbf{A})$. Also, then $\dim(\mathcal{N}(\mathbf{A})) = n - k$. Remember that for the null space $\mathbf{A}\mathbf{v} = 0 \cdot \mathbf{v}$ so its dimension is equal to the number of zero eigenvalues of the symmetric matrix \mathbf{A} .

A.8 Why do orthogonal matrices represent rotations?

TODO: soon, more urgent

<https://math.stackexchange.com/questions/612936/why-are-orthogonal-matrices-generalizations-of-rotation>

<https://www.quora.com/Why-do-orthogonal-matrices-represent-rotations/answer/Julius-Bier-Kirkegaard>

A.9 Matlab program to compress image by using SVD

```
A = im2double(imread('cameraman.bmp'));
[m, n] = size(A);
ranks = [1 2 10 20 round(1/4 * m*n/(m+n+1)) round(1/2 * m*n/(m+n+1))];
cd 'output';
fname = fopen('svd_results.txt','w');
fprintf(fname,'rank, number of elements required, compression ratio, norm2(A-A_k), sigma_(k+1)\n');
for r = ranks
    [U,S,V] = svd(A);
    A_k = U(:,1:r) * S(1:r, 1:r) * V(:, 1:r)';
    % numbers required to stores the approximation
    % size of U + no non-zero singular values + size of V
    numbers_req = numel(U(:,1:r)) + nnz(S(1:r, 1:r)) + numel(V(:,1:r));
    figure,imshow(A_k);
    imwrite(A_k, strcat('rank',int2str(r),'.bmp'));
    % for norm 2 we could do max(max(abs(A - A_k))
    fprintf(fname, strcat(int2str(r), ',', int2str(numbers_req), ',', ...
        num2str((m*r + r + r*n)/(m*n)), ',', num2str(norm(A-A_k,2)), ',', ...
        , num2str(S(r+1,r+1)), '\n'));
end
fclose(fname);
close all; cd '..';
```

The file generated storing the results is listed below. Note how the result confirm the Eckart-Young-Mirsky theorem 5.12.

```
rank, number of elements required, compression ratio, norm2(A-A_k), sigma_(k+1)
1,513,0.0078278,28.3839,28.3839
2,1026,0.015656,21.5064,21.5064
10,5130,0.078278,6.0213,6.0213
20,10260,0.15656,3.5956,3.5956
32,16416,0.25049,2.379,2.379
64,32832,0.50098,1.1571,1.1571
```

A.10 Linear least squares solution

Let's say that we suppose a body is moving roughly in a straight line $y = bx + a$ have measured its position at various points $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$. Ideally, we'd want $y_i = bx_i + a \ \forall i: 1 \leq i \leq n$ to be true. Unless vectors $(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)$ are all l.i., $a + bx_i = y_i, 1 \leq i \leq k$ is impossible to solve as we have more

equations than unknowns, or equivalently the system $\mathbf{A}_{m \times n} \mathbf{x} = \mathbf{b}$, $\mathbf{A} = \begin{bmatrix} 1 & x_1 \\ \dots & \dots \\ 1 & x_k \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ \dots \\ y_k \end{bmatrix}$ is

overdetermined. That follows from the fact that the rows of \mathbf{A} are l.i. and that $m = k > n = 2$.

To summarise, when $\mathbf{A}_{m \times n} \mathbf{u} = \mathbf{b}$, $m > n$ and we have to solve the "tall and skinny" system

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

, where \mathbf{A} has more than n l.i. rows, then \mathbf{A}^{-1} does not exist and there are no solutions. It is impossible to write $\mathbf{b} \in \mathbb{R}^m$ as a linear combination of the columns of \mathbf{A} . However, the column space of \mathbf{A} is in \mathbb{R}^2 . Instead, we want to find the best approximation of \mathbf{b} in \mathbb{R}^2 .

EXAMPLE A.3. ref: <https://www.cs.auckland.ac.nz/courses/compsci369s1c/lectures/GG-notes/CS369-LeastSquares.pdf>

(Linear regression in statistics) Try to fit the measurements $(0, 1), (1, 9), (3, 9), (4, 21)$ in a straight line.

Solution:

Let the straight line be $b_i = u_1 + u_2 x_i$, then we'd have to solve

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 9 \\ 21 \end{bmatrix}$$

The above equation does not have a solution as $\begin{bmatrix} 1 & 9 & 9 & 21 \end{bmatrix}^T$ cannot be written as a linear combination of $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$, $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$, $\begin{bmatrix} 1 & 3 \end{bmatrix}^T$, $\begin{bmatrix} 1 & 4 \end{bmatrix}^T$. **TODO: insert graph here** ■

<https://www.cs.auckland.ac.nz/courses/compsci369s1c/lectures/GG-notes/CS369-LeastSquares.pdf>

The way we work is find a straight line $y = u_1 + u_2 x$ such that the magnitude of the error vector $\mathbf{e} = \begin{bmatrix} e_1 & \dots & e_m \end{bmatrix} = \begin{bmatrix} y_1 - u_1 - u_2 x_1 & \dots & y_m - u_1 - u_2 x_m \end{bmatrix}$, $\|\mathbf{e}\| = e_1^2 + \dots + e_m^2 = (y_1 - u_1 - u_2 x_1)^2 + \dots + (y_m - u_1 - u_2 x_m)^2$ is minimised (hence the "least squares"). In general, when solving the overdetermined system $\mathbf{A}\mathbf{u} = \mathbf{b}$, the error vector is given by

$$\mathbf{e} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}} \quad (\text{A.7})$$

, where $\bar{\mathbf{x}}$ minimises the magnitude of \mathbf{e} . We will prove the following theorem for the least squares solution.

THEOREM A.1. Let \mathbf{A} be an $m \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^m$. Then $\mathbf{A}\mathbf{x} = \mathbf{b}$ always has at least one least squares solution $\bar{\mathbf{x}}$. The following properties hold:

1. $\bar{\mathbf{x}}$ is a least squares solution of $\mathbf{A}\bar{\mathbf{x}} = \mathbf{b}$ iff $\bar{\mathbf{x}}$ is a solution of the normal equations

$$(\mathbf{A}^T \mathbf{A}) \bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b} \quad (\text{A.8})$$

2. \mathbf{A} has linearly independent columns if and only if $\mathbf{A}^T \mathbf{A}$ is invertible. In this case, the least squares solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$ is unique and is given by

$$\bar{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (\text{A.9})$$

Proof. To prove the theorem, we suppose we want to a vector in \mathbb{R}^3 in a 2D plane but what we'll prove holds in general. We want to approximate the \mathbf{b} vector as $\mathbf{A}\bar{\mathbf{x}}$ in the $\mathbf{A}\mathbf{x}$ plane, i.e. in the column space of \mathbf{A} such that the magnitude of $\mathbf{e} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}}$ is minimised. In the figure below one can see (Pythagoras) the best approximation of \mathbf{b} is its projection $\mathbf{p} = \text{proj}_{\mathcal{R}(\mathbf{A})} \mathbf{b} = \mathbf{A}\mathbf{u}$.

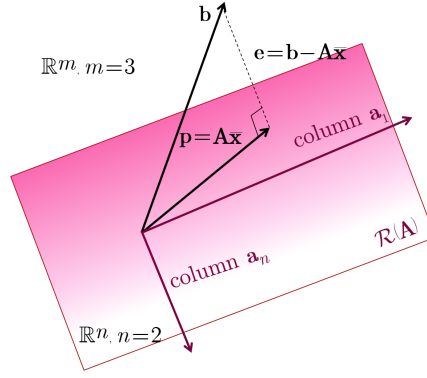


Figure 14: Approximating a 3D vectors in a 2D space by using its projection.

To minimise the length of \mathbf{e} , it must be perpendicular to the column space $\mathcal{R}(\mathbf{A})$, therefore ("fundamental theorem of linear algebra") it must lie in the left nullspace of \mathbf{A} , $\mathcal{N}(\mathbf{A}^T)$, i.e. such that $\mathbf{A}^T \mathbf{e} = 0$. Alternatively, writing \mathbf{A} in columns as $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n]$, we could express the orthogonalirlt as the dot product constraints

$$\begin{aligned} \mathbf{a}_1^T \mathbf{e} &= 0 \\ &\dots \\ \mathbf{a}_n^T \mathbf{e} &= 0 \Rightarrow \\ \mathbf{A}^T \mathbf{e} &= 0 \end{aligned}$$

Either way, we end up with:

$$\begin{aligned} \mathbf{A}^T \mathbf{e} &= 0 \Rightarrow \\ \mathbf{A}^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) &= 0 \Rightarrow \\ \mathbf{A}^T \mathbf{A}\bar{\mathbf{x}} &= \mathbf{A}^T \mathbf{b} \end{aligned}$$

Hence the normal equations are derived. □

Proof. (Part 2, \Rightarrow) We want to know when $\mathbf{A}^T \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$ can be solved for $\bar{\mathbf{x}}$, i.e. when $\mathbf{A}^T \mathbf{A}$ is invertible. Suppose \mathbf{A} has l.i. columns. We want to show that $\mathbf{A}^T \mathbf{A}$ is invertible, or equivalently that the nullspace of $\mathbf{A}^T \mathbf{A}$ is $\mathbf{0}$, i.e suppose

$$\mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{0} \tag{1}$$

Then we want to show that Eq. (1) holds only for $\mathbf{x} = \mathbf{0}$. $\mathbf{A}\mathbf{x} \in \mathcal{R}(\mathbf{A})$ therefore it is orthogonal to the left nullspace of \mathbf{A} . From Eq. (1) we also have that $\mathbf{A}\bar{\mathbf{x}}$ is in the left null space of \mathbf{A} . These two conditions can only hold if the vector is zero:

$$\mathbf{A}\mathbf{x} = \mathbf{0}$$

Since \mathbf{A} 's columns are l.i. it is invertible therefore $\mathbf{x} = \mathbf{0}$. □

Proof. (Part 2, \Leftarrow) Suppose $\mathbf{A}^T \mathbf{A}$ is invertible, then we want to show that \mathbf{A} is invertible too. We first show that the null spaces of $\mathbf{A}^T \mathbf{A}$ and \mathbf{A} are equal (left as exercise) therefore:

$$\begin{aligned} \mathcal{N}(\mathbf{A}^T \mathbf{A}) &= \mathcal{N}(\mathbf{A}) \Rightarrow \\ \dim(\mathcal{N}(\mathbf{A}^T \mathbf{A})) &= \dim(\mathcal{N}(\mathbf{A})) \Rightarrow \\ \text{rank}(\mathcal{N}(\mathbf{A}^T \mathbf{A})) &= \text{rank}(\mathcal{N}(\mathbf{A})) \end{aligned}$$

□

The projection of \mathbf{x} in the column space of \mathbf{A} is

$$\mathbf{p} = \mathbf{A}\bar{\mathbf{x}} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \mathbf{b}$$

Matrix $\mathbf{A}\bar{\mathbf{x}} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T$ is called projection matrix. $\mathbf{Ax} = \mathbf{b}$ has no solution but $\mathbf{Ax} = \mathbf{p}$ has a solution $\bar{\mathbf{x}}$.

COROLLARY A.7.

1. *The projection matrix $\mathbf{P} = \mathbf{A}\bar{\mathbf{x}} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \mathbf{b}$ is symmetric.*
2. *$\mathbf{P}^2 = \mathbf{P}$, as repeated projections give the same result.*
3. *\mathbf{P} is $m \times m$ but of rank only n .*

A.11 Pseudoinverse with QR factorisation

TODO!

<https://see.stanford.edu/materials/lsoeldsee263/05-ls.pdf>

References

- [1] D. Lay, *Linear algebra and its applications*. Addison-Wesley, 2014, pp. 401–424.
- [2] meelo. (). Drawing ellipse from eigenvalue-eigenvector. version 2015-12-04, [Online]. Available: <https://math.stackexchange.com/q/1558345>.
- [3] C.-P. Chen, *Positive definite matrix*, Oct. 2012.
- [4] A. Csaki, “A concise proof of sylvester’s theorem,” 1970. [Online]. Available: <https://pp.bme.hu/ee/article/download/5099/4204>.
- [5] G. Giorgio, “Various proofs of the sylvester criterion for quadratic forms,” vol. 9, p. 55, Oct. 2017. [Online]. Available: https://www.researchgate.net/publication/320661702_Various_Proofs_of_the_Sylvester_Criterion_for_Quadratic_Forms/fulltext/59f366acaca272607e290f09/32066.
- [6] D. Zwick, *Lecture 33 : Positive definite matrices*, 2012.
- [7] *Lecture: Quadratic optimization*, 2017.
- [8] P. Olver, *Numerical analysis lecture notes: 12. minimization*.
- [9] A. Beck, *Introduction to nonlinear optimization*. Society for Industrial and Applied Mathematics, Oct. 2014, ch. 2. DOI: 10.1137/1.9781611973655. [Online]. Available: https://my.siam.org/books/mo19/MO19_ch2.pdf.
- [10] P. O. M. Notes. (). Relative minimums and maximums, [Online]. Available: <http://tutorial.math.lamar.edu/Classes/CalcIII/RelativeExtrema.aspx> (visited on 07/24/2018).
- [11] Y.-F. Wang, *Cs291a: Quadratic programming*, 2017.
- [12] J. R. Bunch and L. Kaufman, “A computational method for the indefinite quadratic programming problem,” *Linear Algebra and its Applications*, vol. 34, pp. 341–370, 1980, ISSN: 0024-3795. DOI: [https://doi.org/10.1016/0024-3795\(80\)90172-X](https://doi.org/10.1016/0024-3795(80)90172-X). [Online]. Available: <http://www.sciencedirect.com/science/article/pii/002437958090172X>.
- [13] A. Sasane and K. Svanberg, *Optimization(sf1811/sf1831/sf1841)*.
- [14] H. Bast, *Seminar: Advanced topics in information retrieval, an existence proof for the singular value decomposition*, Oct. 2004.
- [15] C. Bingham, *Singular value decomposition and digital image compression*, Dec. 2016. [Online]. Available: <https://mse.redwoods.edu/darnold/math45/laproy/fall2016/ChristopherBingham/paper.pdf>.
- [16] T. R. Ĝ. Valiant, *Cs168: The modern algorithmic toolbox lecture 9: The singular value decomposition (svd) and low-rank matrix approximations*, 2017. [Online]. Available: <http://theory.stanford.edu/~tim/s17/l19.pdf>.
- [17] amd. (). A rank-one matrix is the product of two vectors. URL:<https://math.stackexchange.com/q/1545147> (version: 2015-11-25), [Online]. Available: <https://math.stackexchange.com/q/1545147>.
- [18] R. van de Geijn, *Notes on the singular value decomposition*, Sep. 2014. [Online]. Available: <http://www.cs.utexas.edu/users/flame/Notes/NotesOnSVD.pdf>.
- [19] G. Strang, *Introduction to linear algebra*, Fourth. Wellesley, MA: Wellesley-Cambridge Press, 2009, ISBN: 9780980232714 0980232716 9780980232721 0980232724 9788175968110 8175968117.
- [20] R. (<https://math.stackexchange.com/users/326045/rss>), *Transform a sphere to an ellipse using a matrix*, Mathematics Stack Exchange, URL:<https://math.stackexchange.com/q/1948989> (version: 2016-10-01). eprint: <https://math.stackexchange.com/q/1948989>. [Online]. Available: <https://math.stackexchange.com/q/1948989>.
- [21] D. Kalman, *A singularly valuable decomposition: The svd of a matrix*, Feb. 2002. [Online]. Available: [https://datajobs.com/data-science-repo/SVD-\[Dan-Kalman\].pdf](https://datajobs.com/data-science-repo/SVD-[Dan-Kalman].pdf).
- [22] S. Hyde, *Orthonormal bases for the four subspaces using the singular value decomposition*, Feb. 2017. [Online]. Available: <http://jekyll.math.byuh.edu/courses/m343/handouts/svdbasis.pdf>.
- [23] W. Gander, *The singular value decomposition*, Dec. 2008. [Online]. Available: https://www2.math.ethz.ch/education/bachelor/lectures/hs2014/other/linalg_INFK/svdneu.pdf.
- [24] M. Embree, *Singular value decomposition*, Apr. 2017. [Online]. Available: <http://www.math.vt.edu/people/embree/math5524/chapter3.pdf>.
- [25] B. Wang, *Lecture 18: The svd: Examples, norms, fundamental subspaces, compression*, 2016. [Online]. Available: <http://www.sci.utah.edu/~beiwang/teaching/cs6210-fall-2016/lecture18.pdf>.
- [26] V. Guruswami, *Singular value decomposition (svd)*, 2012. [Online]. Available: <http://www.cs.cmu.edu/~venkatg/teaching/CStheory-infoage/SVD-chap.pdf>.
- [27] J. Liu. (). version 2015-11-25, [Online]. Available: <https://www.quora.com/I-was-using-SVD-for-image-compression-But-I-found-out-that-this-is-not-the-technique-that-is-used-in-general-for-this-purpose-ex-JPEG-compression-Why-is-it-so-What-are-some-of-the-disadvantages-of-using-SVD-for-image-compression/answer/Jason-Liu-21>.
- [28] S. P. (<https://math.stackexchange.com/users/166694/stefan-perko>), *Properties of the pseudoinverse of a matrix*, Mathematics Stack Exchange, URL:<https://math.stackexchange.com/q/1496487> (version: 2015-

- 10-25). eprint: <https://math.stackexchange.com/q/1496487>. [Online]. Available: <https://math.stackexchange.com/q/1496487>.
- [29] R. (https://math.stackexchange.com/users/367/rasmus), *Pseudoinverse matrix and svd*, Mathematics Stack Exchange, URL:https://math.stackexchange.com/q/19954 (version: 2011-02-01). eprint: <https://math.stackexchange.com/q/19954>. [Online]. Available: <https://math.stackexchange.com/q/19954>.
- [30] G. Gimel'farb. [Online]. Available: <https://www.cs.auckland.ac.nz/courses/compsci369s1c/lectures/GG-notes/CS369-LeastSquares.pdf>.
- [31] D. Gillies and M. Deisenroth, *Lecture 14: Dimensionality reduction with pca*. [Online]. Available: <https://www.doc.ic.ac.uk/~dfg/ProbabilisticInference/IDAPISlides14.pdf>.
- [32] M.-P. Tsui, *Lecture 23:cholesky factorization*, Mar. 2008.
- [33] D. Leykekhman, *Lecture 5. solving linear systems 3*, 2008.
- [34] P. O. M. Notes, *Directional derivatives*. [Online]. Available: http://tutorial.math.lamar.edu/Classes/CalcIII/DirectionalDeriv.aspx#PD_DirectDeriv_Ex1a.
- [35] D. Auroux, *Gradient:proof that it is perpendicular to level curves and surfaces*, 2010. [Online]. Available: https://ocw.mit.edu/courses/mathematics/18-02sc-multivariable-calculus-fall-2010/2.-partial-derivatives/part-b-chain-rule-gradient-and-directional-derivatives/session-36-proof/MIT18_02SC_notes_19.pdf.
- [36] (). Taylor series expansion, [Online]. Available: <http://fourier.eng.hmc.edu/e176/lectures/NM/node45.html>.
- [37] J. Burkardt, *Vector form of taylor's series, integration in higher dimensions,and green's theorems*, 2014. [Online]. Available: https://people.sc.fsu.edu/~jburkardt/classes/gateway_2014/lecture_week14.pdf.
- [38] *Why/how does the determinant of the hessian matrix, combined with the 2nd derivatives, tell us max., min., saddle points? reasoning behind it?* Mathematics Stack Exchange. [Online]. Available: <https://math.stackexchange.com/q/1985915>.
- [39] N. Jones, *Ee2 mathematics : Functions of multiple variables*. [Online]. Available: <http://wwwf.imperial.ac.uk/~nsjones/ee2mafmv2.pdf>.