Linear Algebra

Iran University of Science and Technology

Dimensionality Reduction Using PCA

By Mohammad Soroush Masoum Beigi, Tara Tajdani

[A Very Short Note on Isomorphism]

V is isomorphic to K^n when dim(V) = n

V is an n-dimensional vector space over field F. Assume B = {v1, v2, ..., vn} is an ordered basis for V, and F^n represents the vector space consisting of all n-tuples (α1, α2, ..., αn) where αi belongs to F. Our goal is to demonstrate that V is isomorphic to F^n.

To prove this, we define a mapping T : V -> F^n as follows: for any v in V, we can express it uniquely as v = α1v1 + α2v2 + ... + αnvn, where α1, α2, ..., αn are elements of F. We define T(v) = (α1, α2, ..., αn), which belongs to F^n.

I) T is well-defined: Since for a given v in V, the coefficients α1, α2, ..., αn are unique, the resulting tuple (α1, α2, ..., αn) is unique as well. Therefore, T(v) is well-defined for any v in V.

II) T is linear: Let x and y be vectors in V, and let α and β be elements of F. We can express x and y as x = a1v1 + a2v2 + ... + anvn and y = b1v1 + b2v2 + ... + bnvn, where ai and bi are unique elements of F. T(x) = (a1, a2, ..., an) , T(y) = (b1, b2, ..., bn) Now, consider αx + βy: αx + βy = (αa1v1 + αa2v2 + ... + αanvn) + (βb1v1 + βb2v2 + ... + βbnvn) = (αa1 + βb1)v1 + (αa2 + βb2)v2 + ... + (αan + βbn)vn Hence, T(αx + βy) = (αa1 + βb1, αa2 + βb2, ..., αan + βbn) = (αa1, αa2, ..., αan) + (βb1, βb2, ..., βbn) = α(a1, a2, ..., an) + β(b1, b2, ..., bn) = αT(x) + βT(y) Therefore, T is a linear transformation.

III) T is one-to-one: If T(x) = T(y), then T(a1v1 + a2v2 + ... + anvn) = T(b1v1 + b2v2 + ... + bnvn), implying (a1, a2, ..., an) = (b1, b2, ..., bn). Consequently, a1 = b1, a2 = b2, ..., an = bn, which in turn leads to a1v1 + a2v2 + ... + anvn = b1v1 + b2v2 + ... + bnvn. Therefore, T is a one-to-one mapping, and it follows that x = y.

IV) T is onto: Let (α1, α2, ..., αn) be an element of F^n, where α1, α2, ..., αn belong to F. We can construct v = α1v1 + α2v2 + ... + αnvn, which belongs to V. As a result, T(v) = (α1, α2, ..., αn).

Find an isomorphism α : V → K^n where ker(α) = {0}

To find an isomorphism α : V → K^n such that the kernel (ker(α)) is {0}, we must first understand the meaning of these terms.

The kernel of a linear transformation $\alpha : V \to W$ refers to the set of all vectors in V that are mapped to the zero vector in W. In simpler terms, $\ker(\alpha) = \{v \in V \mid \alpha(v) = 0\}$.

To obtain an isomorphism $\alpha : V \to K^n$ with $\ker(\alpha) = \{0\}$, we need to locate a linear transformation α that is both one-to-one and onto, while ensuring its kernel is {0}. One approach is to select a basis {v1, v2, ..., vn} for V and define α as follows:

$\alpha(v1) = (1, 0, \ldots, 0)$ $\alpha(v2) = (0, 1, \ldots, 0)$ ... $\alpha(vn) = (0, 0, \ldots, 1)$

Essentially, we map each basis vector of V to a distinct standard basis vector of K^n. This definition establishes a linear transformation $\alpha : V \to K^n$.

To demonstrate that α is one-to-one, we can prove that $\ker(\alpha) = \{0\}$. Let's assume $\alpha(v) = 0$ for some $v \in V$. We can express v as a linear combination of the basis vectors:

$v = a1v1 + a2v2 + \ldots + anvn$

Since α is linear, we have:

$0 = \alpha(v) = a1\alpha(v1) + a2\alpha(v2) + \ldots + an\alpha(vn)$

As the basis vectors are linearly independent, this implies that $a1 = a2 = \ldots = an = 0$. Consequently, $v = 0$, leading to $\ker(\alpha) = \{0\}$.

To demonstrate that α is onto, we can prove that its range spans K^n. Let (x1, x2, ..., xn) be an arbitrary vector in K^n. We can define a vector v in V as follows:

$v = x1v1 + x2v2 + \ldots + xnvn$

Since α is linear, we have:

$\alpha(v) = x1\alpha(v1) + x2\alpha(v2) + \ldots + xn\alpha(vn) = (x1, x2, \ldots, xn)$

Thus, α is onto.

In conclusion, α is a one-to-one and onto linear transformation with $\ker(\alpha) = \{0\}$, signifying that it is an isomorphism from V to K^n.

Is α a Homomorphism?

Yes, α is a homomorphism due to its nature as a linear transformation connecting vector spaces over the same field.

Let's recall that a homomorphism between two vector spaces V and W over a field K is a linear transformation f: V → W that maintains the structure of the vector spaces. In other words, it adheres to the following properties:

1. f(u + v) = f(u) + f(v) for all u, v ∈ V
2. f(cu) = cf(u) for all u ∈ V and c ∈ K

In the specific scenario of α : V → K^n, we observe:

1. α(u + v) = α(u) + α(v) for all u, v ∈ V. This property holds true because α is defined as a linear combination of the basis vectors of V.
2. α(cu) = cα(u) for all u ∈ V and c ∈ K. This property holds true as α is a linear transformation.

Consequently, α satisfies the requirements of a homomorphism, establishing it as a homomorphism from V to K^n.

What can one say about Isomorphisms, studying only the properties of kernels?

The kernel of an isomorphism, a concept from abstract algebra, is closely related to the kernel of a homomorphism.

An isomorphism refers to a bijective homomorphism between two algebraic structures, preserving their respective structures. In the context of an isomorphism, the kernel represents the set of elements in the domain that are mapped to the identity element in the codomain.

To provide a more precise definition, let's consider an isomorphism f: G → H between two groups G and H. The kernel of f is defined as the set of elements in G that map to the identity element in H:

kernel(f) = {g ∈ G | f(g) = e_H}

Here, e_H denotes the identity element in H. The kernel of an isomorphism always serves as a normal subgroup of the domain group and is a subgroup of the domain group itself. In fact, it is the largest normal subgroup of the domain group that, through the isomorphism, is mapped to the identity element in the codomain group.

The kernel holds significance in group theory, enabling the study of group structures and their subgroups.

To establish the proof that the kernel of an isomorphism solely consists of the identity element, we must first comprehend the meaning of an isomorphism and the kernel of a function.

An isomorphism signifies a bijective function linking two algebraic structures while preserving their respective structures. In simpler terms, it maps elements from one structure to another, ensuring the operations and relationships among the elements remain unchanged.

The kernel of a function refers to the set of elements within the domain that are mapped to the identity element in the codomain. Essentially, it encompasses the elements that are transformed to zero or the neutral element of the operation.

Assuming we have an isomorphism between two algebraic structures A and B, represented by the functions f: A -> B and g: B -> A, we aim to demonstrate that the kernel of f solely contains the identity element of A.

Let's assume the existence of an element 'a' in the kernel of f, such that 'a' is not the identity element of A. This implies that f(a) = e, where 'e' represents the identity element of B.

Now, let's apply the function g to both sides of the equation: g(f(a)) = g(e). As g serves as the inverse of f, we have g(f(a)) = a. Consequently, we obtain a = g(e).

However, since 'e' denotes the identity element of B, we have g(e) = e. Therefore, a = e, contradicting our initial assumption that 'a' is not the identity element of A.

Consequently, we have successfully proven that in an isomorphism, the kernel solely consists of the identity element.

[Recenter Data – Query]

Find the center of the data

Given an arbitrary dataset D with dim = 2, let's assume that each data point is represented as $X_i = (x_i, y_i)$. Now, we can calculate the sample mean $\bar{X}_n$ for this dataset.

The sample mean $\bar{X}_n$ is computed as the average of all the data points in the dataset. Mathematically, it can be expressed as:

$\bar{X}_n = 1/n * (\Sigma X_i)$

Expanding this expression, we have:

$\bar{X}_n = 1/n * (\Sigma x_i, \Sigma y_i)$

This can further be simplified by applying the summation separately for the x-coordinates and y-coordinates:

$\bar{X}_n = (1/n * \Sigma x_i, 1/n * \Sigma y_i)$

In summary, the sample mean $\bar{X}_n$ for the given dataset D, with a dimension of 2, is equal to $(1/n * \Sigma x_i, 1/n * \Sigma y_i)$.

Shift the Center in the way that it lies on the origin O(0, 0)

Our goal is to reposition the center of the dataset to align with the origin O(0, 0). We define C as the center of the dataset, given by $C = (x_0, y_0)$. To achieve this, we perform a straightforward operation of subtracting C from each data point in X_n. This results in a modified dataset, D', where each data point X_i is transformed to $X'_i = (x_i - x_0, y_i - y_0)$. As a result, the center of the modified dataset D' is now located precisely at the origin O(0, 0).

Find a Homomorphism T: V → V such that $T(\bar{X}_n)$ = O(0, 0) and all other points are positioned respectively. Call the relocated dataset D’

In order to shift the dataset X_n to the origin (0,0) while maintaining the relative positions of all other points, we can employ a translation transformation represented by the homomorphism T: V → V.

The center of the dataset Xn is denoted by (x0, y0) and can be calculated as follows: x0 = 1/n (Σ xi) y0 = 1/n (Σ yi)

By defining the homomorphism T: V → V as: T(xi, yi) = (xi - x0, yi - y0)

we achieve the desired shift, where each point (xi, yi) is transformed to (xi - x0, yi - y0). This preserves the relative positions of the points while relocating the dataset's center to the origin (0,0).

Find the center of the data

The kernel of a linear transformation T is the set of vectors in the domain that get mapped to the zero vector in the codomain. In this case, both the domain and codomain are the vector space V, and the zero vector is represented as (0,0).

To determine ker(T), we solve the equation T(x,y) = (0,0). By substituting the definition of T, we obtain:

T(x,y) = (x - x0, y - y0) = (0,0)

This implies that x = x0 and y = y0. Hence, the kernel of T comprises vectors of the form (x0, y0), forming a one-dimensional subspace of V.

In simpler terms, ker(T) can be expressed as {(x0, y0) | x0, y0 ∈ R}.

[Best Fit]

Consider an arbitrary line C = αX

In semi-parametric regression, we don't assume a strictly linear relationship between the variables. Instead, we introduce flexibility by using a smoothing function to estimate the local trend of the response variable. The semi-parametric regression model can be represented as follows: C = βX + f(Z) Here, β represents the linear regression coefficient, X denotes the predictor variable, and f(Z) represents the non-parametric smoothing function that estimates the trend of the response variable based on the values of other predictor variables Z. This model allows for non-linear relationships between the predictor and response variables while still incorporating a linear component through the regression coefficient β.

Show that line C is a subspace of V containing D’

In order to demonstrate that line C, which contains D', is a subspace of V, we need to confirm two conditions:

1. Line C is closed under vector addition.
2. Line C is closed under scalar multiplication.

Let's consider line C = αX, where X is a vector in V and α is a scalar. Since D' is a dataset in V, we can assume that X and all other vectors in C are also in V.

To establish closure under vector addition, we need to demonstrate that for any two vectors u and v in C, their sum u + v is also in C. Suppose u = αX and v = βX, where α and β are scalars. Their sum is:

$u + v = \alpha X + \beta X = (\alpha + \beta)X$

Since α + β is also a scalar, u + v is in C. Hence, C is closed under vector addition.

To establish closure under scalar multiplication, we need to show that for any vector u in C and any scalar k, the product ku is also in C. Let u = αX be an arbitrary vector in C, where α is a scalar. The product ku is:

$ku = k(\alpha X) = (k\alpha)X$

Since kα is also a scalar, ku is in C. Consequently, C is closed under scalar multiplication.

Therefore, we have demonstrated that line C is a subspace of V containing D', as it fulfills the two conditions for a subspace: closure under vector addition and closure under scalar multiplication. Additionally, it encompasses all the vectors in D'.

Project all points on line C using the projective transformation mentioned in the previous section

To perform the projection of all points onto line C using the mentioned projective transformation, we must establish the coordinates of the points on line C within a suitable n-dimensional vector space Πn.

Assuming line C exists within a two-dimensional vector space with coordinates (x, y), we can define the points on line C as: $C = \{(x, y) \mid y = \alpha x\}$

For projecting the points on line C using the earlier projective transformation αj: Πn → Kj, we can define the coordinates of the points within a three-dimensional vector space Π3 as: P = (x, y, 1)

Setting the third coordinate of each point to 1 ensures invertibility of the transformation. Consequently, we can apply the projective transformation αj to the points in Π3 as follows: αj(P) = (x, y, 1) → kj

The resulting transformation of each point is represented by the scalar kj within the scalar field Kj.

In the case of line C = αX, where X is a vector in a two-dimensional vector space, we can define the points on the line as: C = {tX | t ∈ R}

To perform the projection of the points on line C using the projective transformation, we can define the coordinates of the points within a three-dimensional vector space Π3 as: P = (tX, 1)

Then, the projective transformation αj can be applied to the points in Π3 as follows: αj(P) = (tX, 1) → kj

As a result, each point undergoes a transformation resulting in the scalar kj within the scalar field Kj.

Find a basis B for the projective space C

Consider a vector space V with a basis B = {v1, v2, ..., vn}. We can express any vector x in V as a linear combination of the basis vectors. Now, let's focus on the set of lines passing through the origin that can be generated by scaling the basis vectors, resulting in a space denoted as C. To find a basis for the projective space C, we need to identify a set of lines within C that are linearly independent. This can be achieved by selecting one vector from each line in C, denoted as u1 = v1, u2 = v2, ..., un = vn. By assuming the existence of scalars b1, b2, ..., bn such that a linear combination of the chosen vectors equals zero, we can substitute the expressions for ui as linear combinations of the basis vectors. This leads to a system of linear equations that can be solved. If the only solution to this system is b1 = b2 = ... = bn = 0, then the vectors u1, u2, ..., un are linearly independent. Consequently, the set of lines generated by the basis vectors B, represented as {αv | α ∈ R, v ∈ B}, forms a basis for the projective space C.

Show that every vector on line C is of the form Pn 1 tibi where bi ∈ B and ti ∈ F

Let's recall that line C is defined as $C = \alpha X$, where X is a vector in a two-dimensional vector space and α is a scalar. We have a basis $B = \{v1, v2\}$ for this vector space, where v1 and v2 are linearly independent. By expressing vector X in terms of the basis B, we get $X = a1v1 + a2v2$, where a1 and a2 are scalars. Consequently, line C can be represented as $C = \alpha(a1v1 + a2v2) = (\alpha a1)v1 + (\alpha a2)v2$. Let P be a point on line C with coordinates (x, y) in the two-dimensional vector space. We express P as $P = (x, y) = t1(v1) + t2(v2)$, where t1 and t2 are scalars.

Since P lies on line C, we have $(x, y) = \alpha(a1v1 + a2v2) = \alpha(t1v1 + t2v2)$. By equating the components, we find $\alpha t1 = x/a1$ and $\alpha t2 = y/a2$. As α is a nonzero scalar, we can choose it freely, allowing us to solve for t1 and t2. Specifically, we obtain $t1 = x/(\alpha a1)$ and $t2 = y/(\alpha a2)$. Hence, we can express P as $P = (x/(\alpha a1))(v1) + (y/(\alpha a2))(v2) = (x/a1)(1/(\alpha))(v1) + (y/a2)(1/(\alpha))(v2)$. Defining $bi = (1/(\alpha))vi$, where vi is the i-th basis vector in B, we can rewrite P as $P = (x/a1)(bi) + (y/a2)(bi+1) = Pn1\ tibi$, where bi ∈ B and ti ∈ F.

Find a basis for V

To find a basis for the vector space V, we aim to identify a collection of vectors that are both linearly independent and span V. Consider V as the vector space comprising all 2x2 matrices with real entries, satisfying the condition that the sum of entries in each row and column is zero. Let A represent an arbitrary matrix in V with entries aij. One method to obtain a basis for V is by employing the following set of matrices: B = {E11 - E22, E12 + E21, E12 - E21, E11 + E22}, where Eij refers to the 2x2 matrix containing a 1 in the (i,j)-th position and zeroes elsewhere. To establish that the matrices in B meet the requirements of being linearly independent and spanning V, we proceed as follows: Linear independence: Suppose there exist scalars c1, c2, c3, and c4 such that: $c1(E11 - E22) + c2(E12 + E21) + c3(E12 - E21) + c4(E11 + E22) = 0$.

This results in the following equations for the entries of the resultant matrix: $(c1 + c4)a11 + (c2 + c3)a12 = 0$, $(c2 + c3)a21 + (c1 + c4)a22 = 0$, $(c1 - c4)a12 + (c2 - c3)a11 = 0$, $(c2 - c3)a22 + (c1 - c4)a21 = 0$.

Since the sum of entries in each row and column of A is zero, we have $a11 + a22 = -a12 - a21 = 0$. Consequently, the equations simplify to: $(c1 + c4)a11 = 0$, $(c1 + c4)a22 = 0$, $(c1 - c4)a12 = 0$, $(c1 - c4)a21 = 0$.

Assuming A is non-zero, at least one of the entries a11, a22, a12, or a21 is non-zero. Without loss of generality, let's assume a11 is non-zero. Then, the first two equations above yield c1 + c4 = 0, and the third and fourth equations yield c1 - c4 = 0. Solving these equations results in c1 = c4 = 0 and c2 = c3 = 0, indicating that the matrices in B are linearly independent. Spanning: Let A be an arbitrary matrix in V. We can express A as a linear combination of the matrices in B as follows: A = (1/2)(a11 + a22)(E11 + E22) + (1/2)(a21 - a12)(E12 - E21) + (1/2)(a21 + a12)(E12 + E21) + (1/2)(a11 - a22)(E11 - E22).

Hence, the matrices in B span V. Consequently, the set of matrices B = {E11 - E22, E12 + E21, E12 - E21, E11 + E22} serves as a basis for the vector space V.

Show that for every two arbitrary vector spaces V1 and V2 with B1 and B2 as basis respectively, if B1 and B2 are homomorphic, then V1 and V2 are homomorphic

Let's suppose that V1 and V2 are two vector spaces with respective bases B1 and B2, and B1 and B2 are homomorphic. Homomorphic means that there exists a bijective linear map f: V1 -> V2, where f(B1) = B2, indicating that the elements of B1 are mapped to the corresponding elements of B2 under f. We now aim to demonstrate that the linear structure of V1 and V2 is maintained by f. In other words, for any vectors v1 and w1 in V1, and any scalars a and b, the following holds: f(av1 + bw1) = af(v1) + bf(w1) To prove this, let v1 and w1 be arbitrary vectors in V1, and let a and b be arbitrary scalars. Since B1 forms a basis for V1, we can express v1 and w1 as linear combinations of the basis elements: v1 = c1b1 + c2b2 + ... + cnbn w1 = d1b1 + d2b2 + ... + dnb Here, c1, c2, ..., cn and d1, d2, ..., dn are scalars, while b1, b2, ..., bn are the elements of B1. By applying the linear map f to the above equations, we obtain:

f(v1) = c1f(b1) + c2f(b2) + ... + cnf(bn) = c1b2 + c2b2 + ... + cnbn f(w1) = d1f(b1) + d2f(b2) + ... + dnf(bn) = d1b2 + d2b2 + ... + dnb Note that we utilized the fact that f(B1) = B2. Now, we can compute f(av1 + bw1) as follows: f(av1 + bw1) = f(ac1b1 + ac2b2 + ... + acnbn + bd1b1 + bd2b2 + ... + bdnb) = af(c1b1 + c2b2 + ... + cnbn) + bf(d1b1 + d2b2 + ... + dnb) = af(v1) + bf(w1) Thus, we have demonstrated that f preserves the linear structure of V1 and V2, implying that V1 and V2 are isomorphic (or homomorphic).

Find an isomorphism from B ’ to B

Let V1 and V2 be vector spaces with bases B1 and B2, respectively. We can establish an isomorphism f: V1 -> V2 as follows:

> Define f: B1 -> B2 by assigning f(b1i) = b2i for i = 1, 2, ..., n, where b1i and b2i represent the i-th basis vectors in B1 and B2, respectively.
> Extend f linearly to all of V1. For any vector v1 in V1, express v1 as a linear combination of the basis vectors in B1, such as v1 = a1b1 + a2b2 + ... + anbn. Then, set f(v1) = a1f(b1) + a2f(b2) + ... + anf(bn). By construction, f is a bijective linear map, making it an isomorphism between V1 and V2. The fact that B1 and B2 are homomorphic implies the existence of a linear map f': V1 -> V2 such that f'(B1) = B2. However, the above construction demonstrates that we can always choose f to be the specific isomorphism that maps each basis vector in B1 to its corresponding basis vector in B2.

Show That V is isomorphic to C

To show that the vector space V is isomorphic to the line C = αX, where α is a scalar and X is a vector, we need to demonstrate the existence of a bijective linear transformation between V and C.

Let's define the linear transformation T: V → C as follows: For any vector v in V, we can express it as a linear combination of some basis vectors, say v = a1v1 + a2v2 + ... + anvn, where ai are scalars and vi are the basis vectors of V. Then, we define T(v) as T(v) = α(a1v1 + a2v2 + ... + anvn) = αv, where α is the scalar associated with the line C.

Now, let's verify that T is a linear transformation. For any vectors u and v in V and any scalar c, we have: T(u + cv) = α[(a1u1 + a2u2 + ... + anun) + c(a1v1 + a2v2 + ... + anvn)] = α[(a1u1 + a1cv1) + (a2u2 + a2cv2) + ... + (anun + ancvn)] = α[a1(u1 + cv1) + a2(u2 + cv2) + ... + an(un + cvn)] = α(u + cv) = T(u) + cT(v).

Therefore, T satisfies the linearity property.

To show that T is bijective, we need to demonstrate both injectivity and surjectivity.

Injectivity: Suppose u and v are any two vectors in V, and T(u) = T(v). Using the definition of T, we have: T(u) = αu = αv = T(v).

Since $\alpha \neq 0$ (as it represents the scalar associated with the line C), we can divide both sides of the equation by α, resulting in $u = v$. Hence, T is injective.

Surjectivity: Let c be any scalar, and consider the vector $w = (1/\alpha)cv$. Since $T((1/\alpha)cv) = \alpha((1/\alpha)cv) = cv$, we can see that T is surjective, as for any vector cv in C, we can find a vector $(1/\alpha)cv$ in V such that $T((1/\alpha)cv) = cv$.

Therefore, T is a bijective linear transformation that maps V to C, establishing the isomorphism between V and C.