x'~f(u,Σ) tal que E[x' ]=u y D(x' )=Σ. Sea x=x'-u.

E[x]=0. Sea z1=c1T x Queremos maximizar Var (Z1 )=c1T Σc1. Observemos que maximizar Var (Z1 ) es lo mismo que maximizar .

Pero . Luego el valor de c1 que maximizar Var (Z1 ) es la solución del sistema Σc1=0. Esto es el espacio nulo de la matriz Σ (En este caso un conjunto infinito).

Observemos que el Hessiano de es  y que bajo el supuesto de que Σ es definida positiva el determinante del Hessiano es positivo.

Busquemos ahora un c1 de norma unitaria. Es decir, tal que

c1T c1=‖C1 ‖2=1. Ahora maximizaremos.

\_- 

Los valores propios y vectores propios de una matriz satisfacen Avi= λi vi.

Los λi son los valores propios y los vi son los vectores propios.

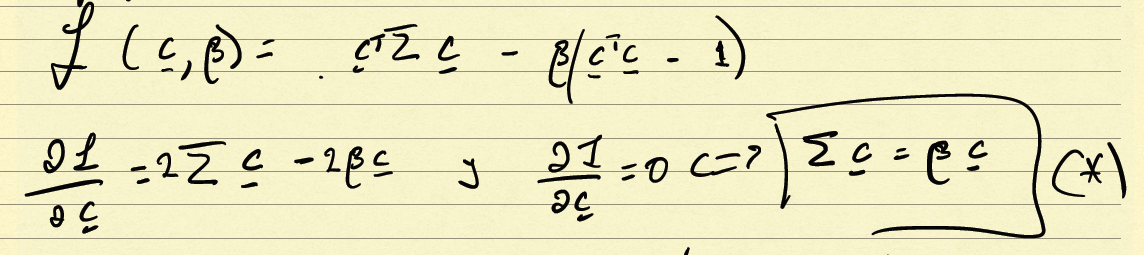
Observemos que si dejamos la función

entonces ∀a∈R se tiene que

Así que sin pérdida de generalidad puede asumirse que ‖C1 ‖=1.

Luego maximizaremos sujeto a ‖C‖2=Σci2=1.

El lagrnagano será:

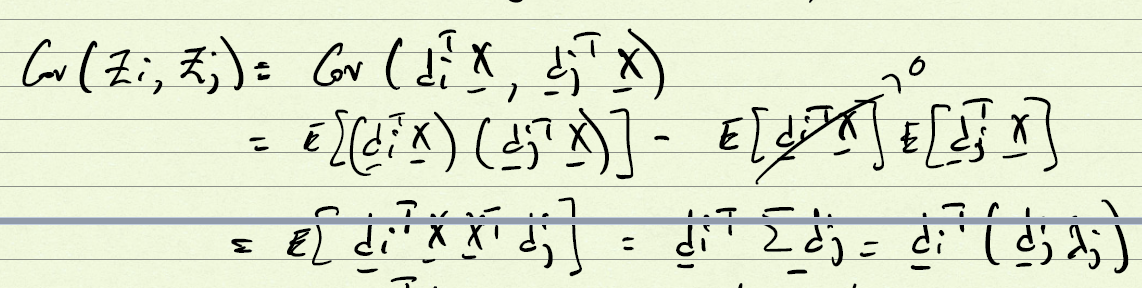


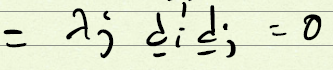
Los valores de c y β que satisfacen esta ecuación son los vectores y valores propios de Σ, digamos d1,d2…dp y λ1,λ2…λp .

Observemos que si di es un vector propio de entonces

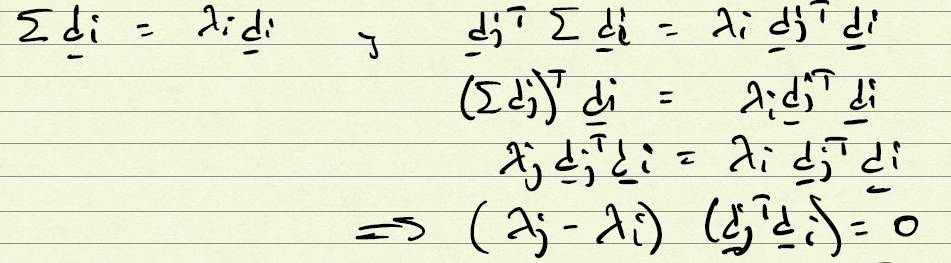
, asumiendo que ‖di ‖=1 entonces se tiene que var (diT di )=λi. Luego debemos escoger di. correspondiente al máximo valor propio.

Sea y -, entonces



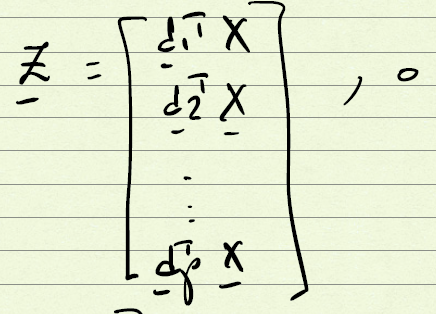


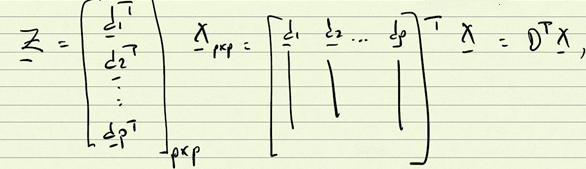
Pero los vectores propios son ortogonales. Para verlo notemos que

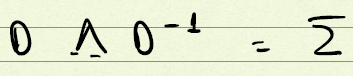


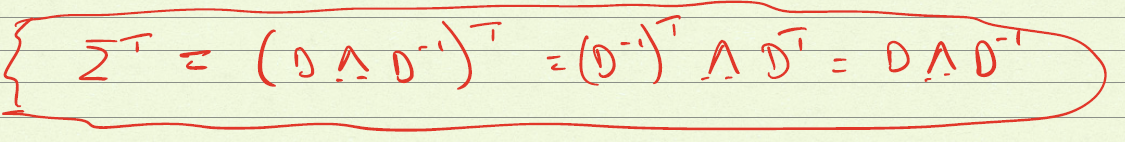
Y esto es cierto para todo si y solo si , pues en general.

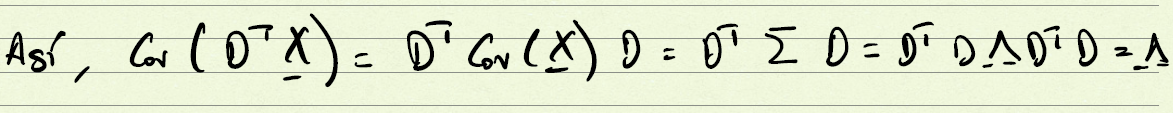
De esta manera definimos

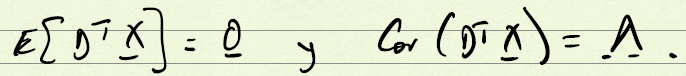


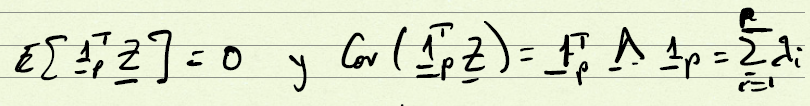


Donde  es la descomposición espectral de . Con es simétrica



Así, 

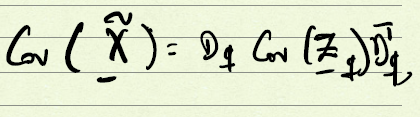
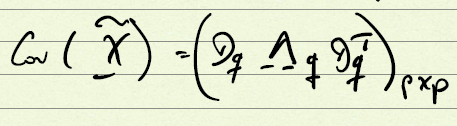
Observemos que 

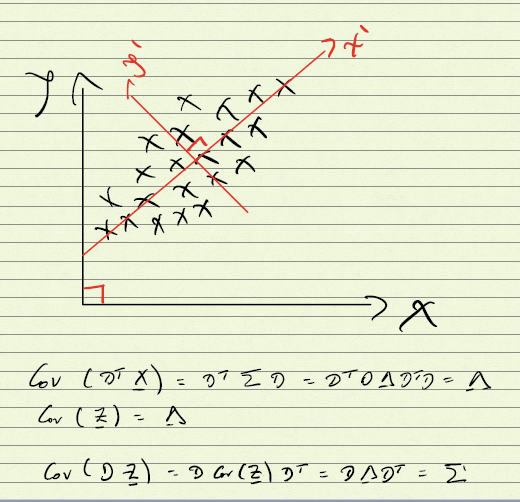
Observemos que .

La varianza de la suma de los es la suma de los

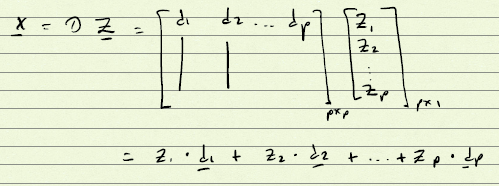
Por otro lado, si entonces

Sea los primeros elementos del vector .

Así,. Dejamos la reconstrucción de como . Observemos que  o sea que 

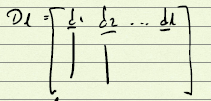


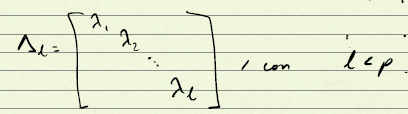
Esto quiere decir que se puede representar usando los componentes principales:



Luego los forman un diccionario para representar a los datos.

Reducción de la dimensionalidad

Consideremos aproximar por donde  donde 

.

Entonces es la representación de en el espacio de primeros componentes principales y es la reconstrucción de usando los primeros de Componentes Principales.

¿Cuál es la calidad de la reconstrucción?

