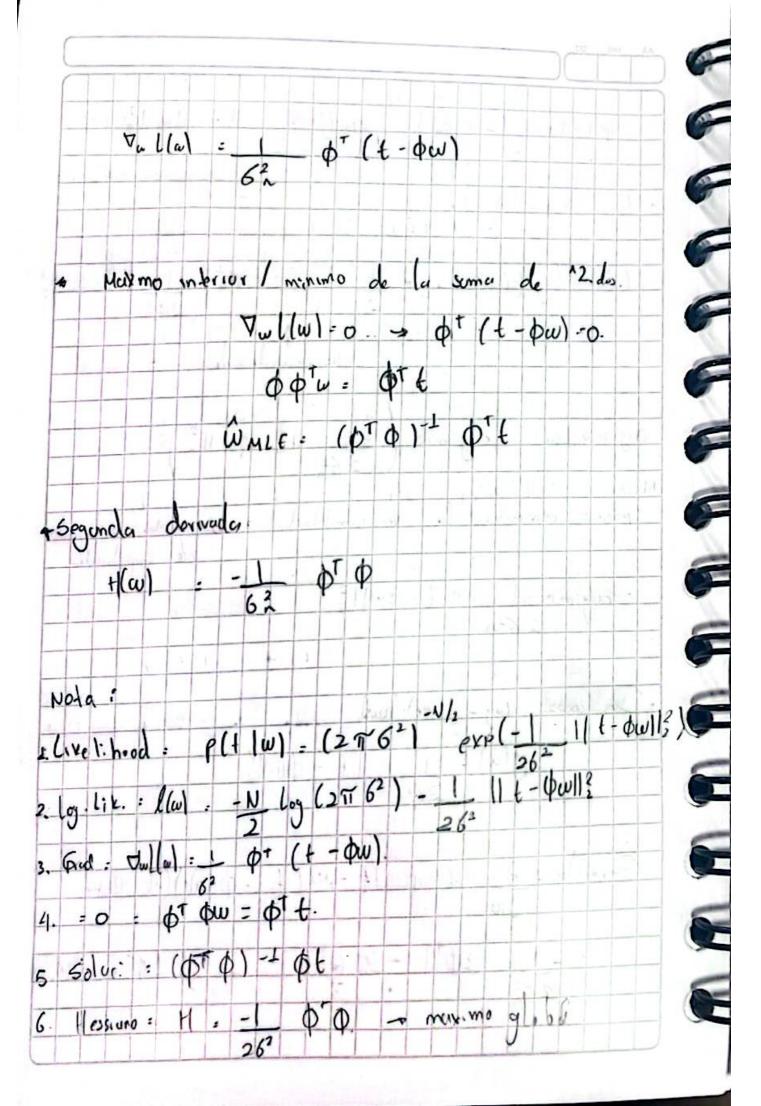
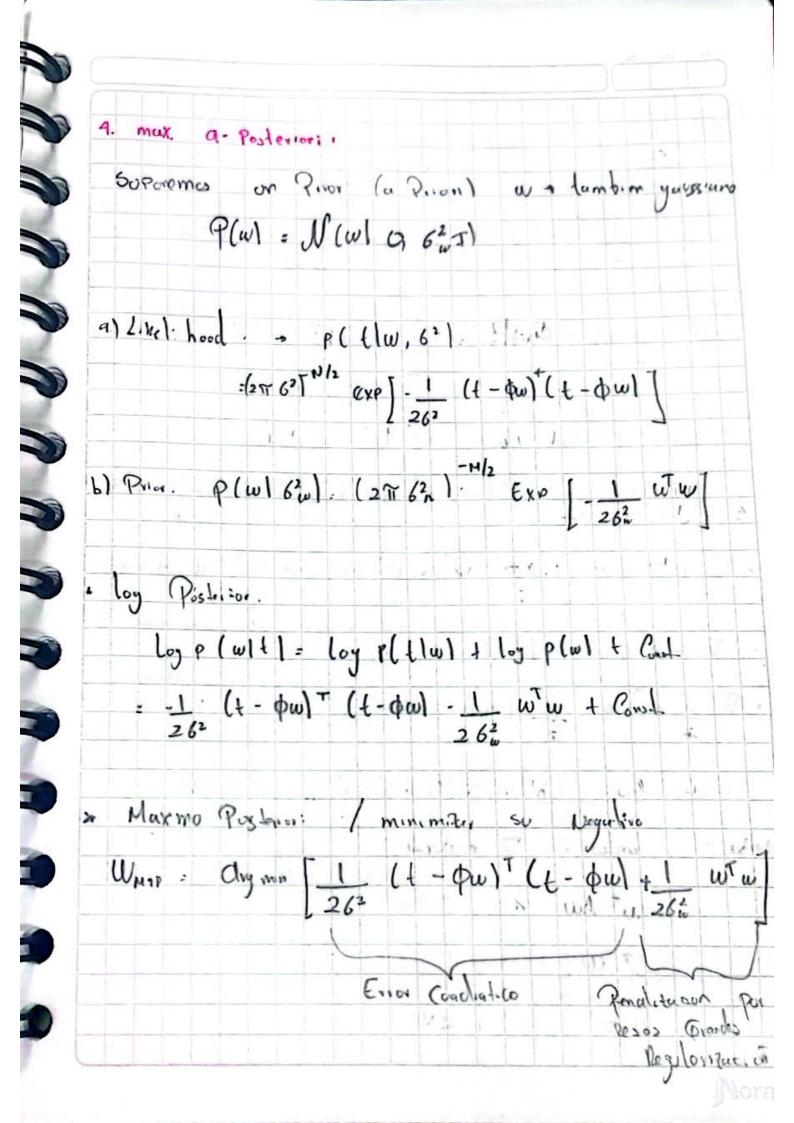
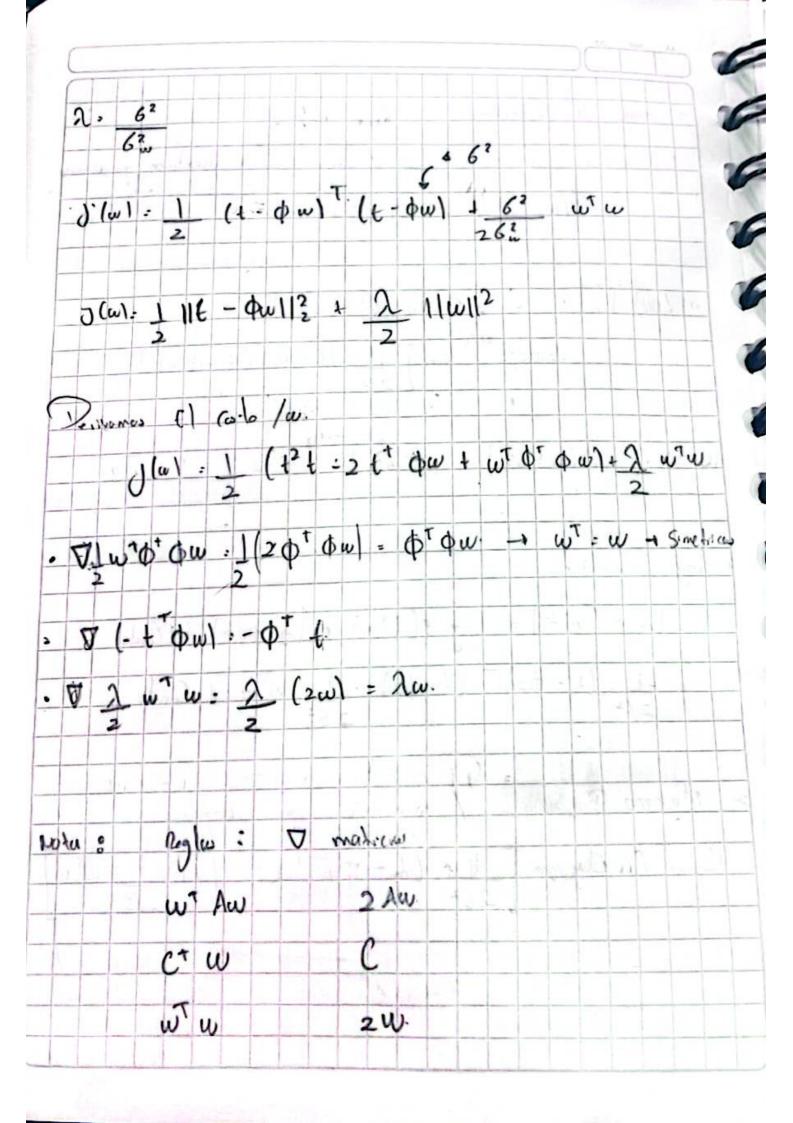
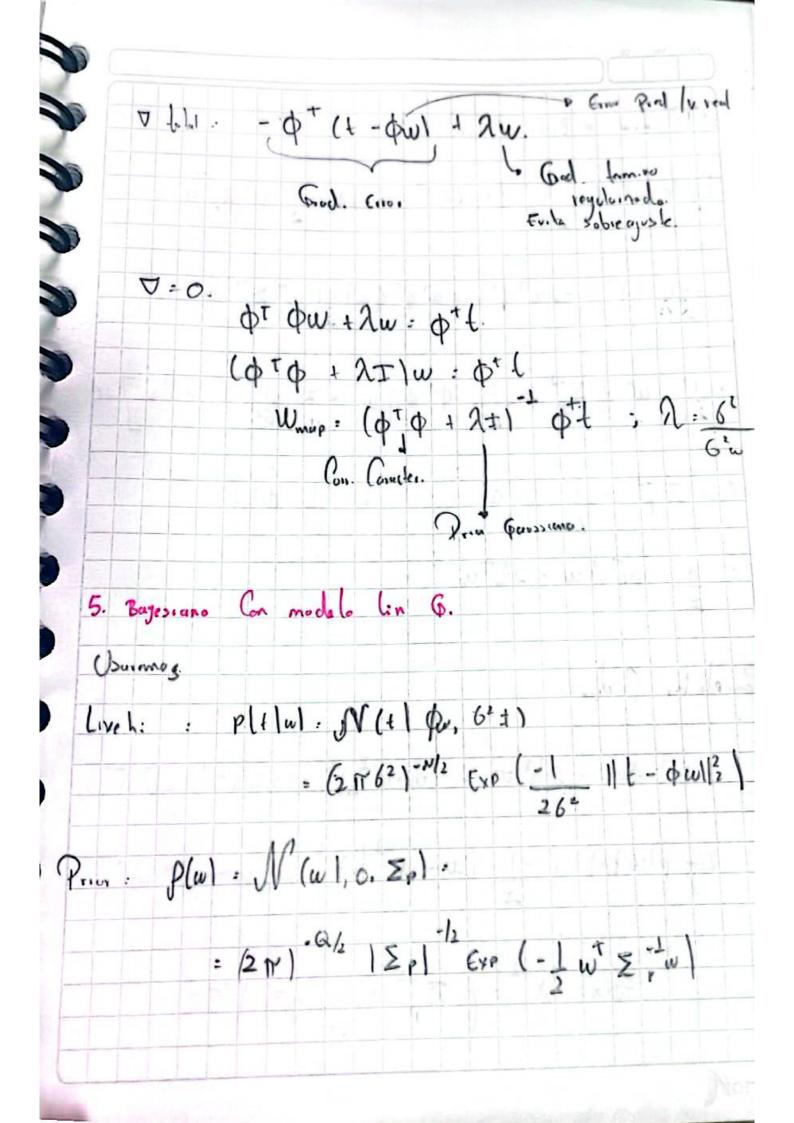


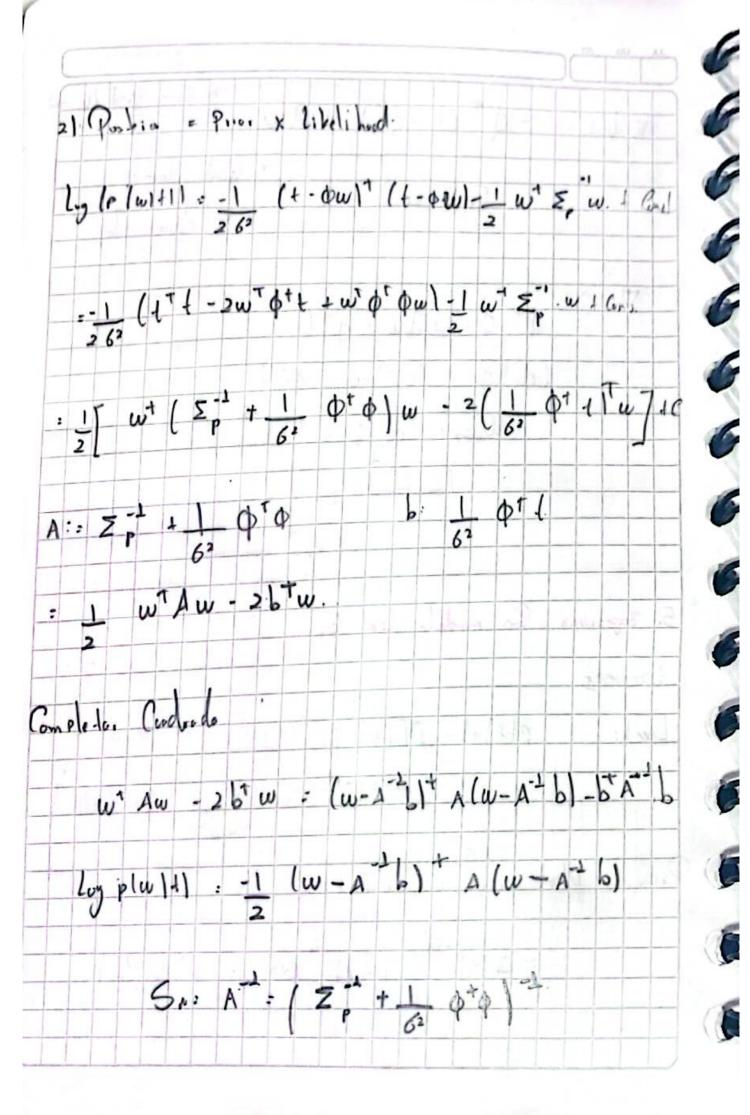
P(flw1: (2 m 62) - W/2 exp(-1 11t- oull?) · Loy - Vongen Hod L(w) ! loy P(tlw). - N loy (12 7 6 2) - 1 11 t- owll2 No depende de Org max 1(w): cry max (-1 11t-qw112) poto : max m. to. (a Ver. similitud, = iminimites le some de cadradas = origina 1 11 t- owl12. · Dorivamos Lug - Likelihood g(w): 1 | t - pw1/2 Dermoda: 2 116- pull? : (t-pu) (t-pu) dw $=\frac{1}{26n^2}$ $(2\phi^{T}\xi - 2\phi\phi^{T}\omega)$ $=\frac{1}{6n^2}\left(\phi^{\dagger}\xi-\phi\phi\right)$

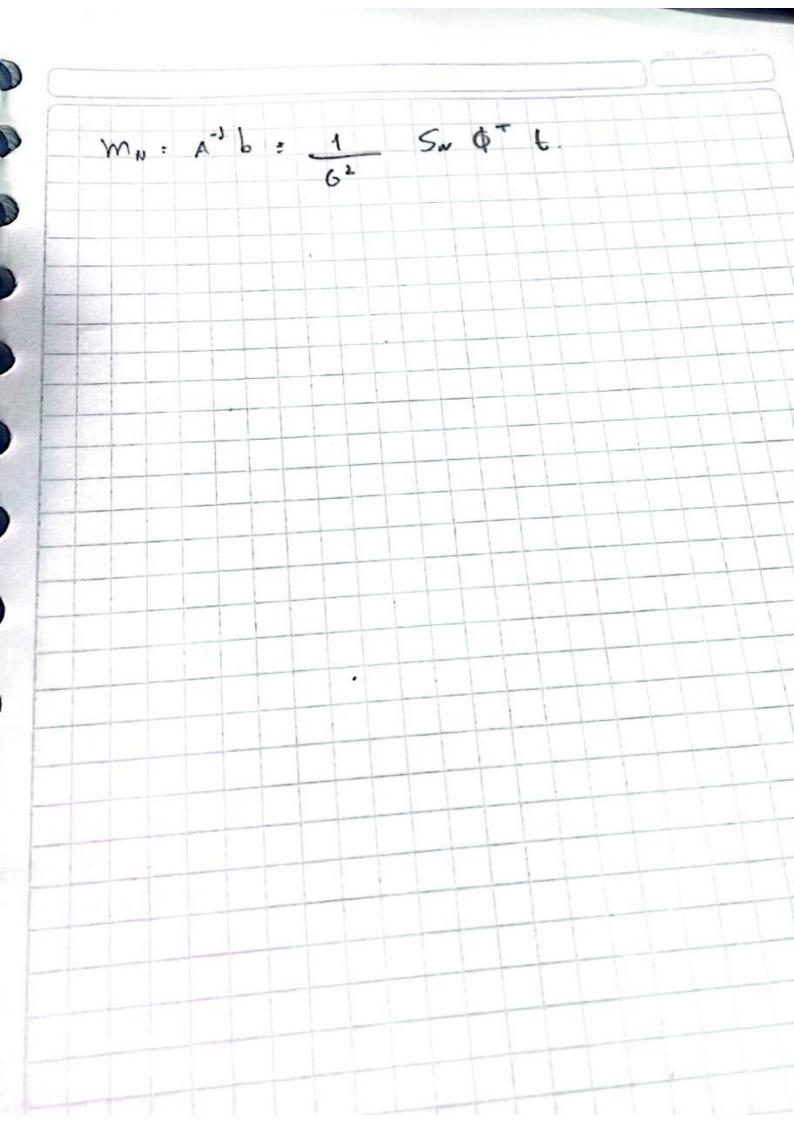












Regression Rigida for Kernel Queremos permity no linealized sin trabaja: explicitamento Con w en 12°, usando un kernel $(x, x') = \phi(x) \phi(x')$ Ridge en espacio de característicos min 11f - 0w12 + 1 1w12 solucion dual (teorema de representación): la solicion ω pertenece al subespecio generato por las filas de Φ, es aterir, existe α ε R" tal que ω=Φ'α. 1-0w=f-00Tx=f-Kx K= DDT ∈ RNAN es la motriz de Kernel Con Kij: k(Xi, X,) El problema en & es: min 11t - Kx112 + Xx PDTX = 11f - Kx112 + Xx Kx Demando e igualando a Cero: (K + XI) x = t => x = (K+ XI) 1 la predicción para un nuevo X. Usa: t. = p(x.) Tw= b(x.) TΦTX = Kx X Dende K. = (K(Xx, X,), ..., K(Xx, XN)) Par & tanto: Ex = K* (K + XI) + Interpretacion: keinel ridge permite operar con k directamente sinco.

Regiosion Poi picceso Gaussiano Un Pieceso Gaussiano (GP) pone un prior directamente sobre Aunciones f(·) ((·)~ GP(0, K(·,·)) fn= ((Xn) + 1/n (on 1/n N 10, 5%) Prior Conjunta sobre labres observados y no observados Sea f=(f(xn), ..., f(xu)) Entences $P(f) = \mathcal{N}(0, K), K_{ij} = K(x_i, x_j)$ t=f+1 con 1~ N(0,02 I) por la tento. P(1) = N(0, k+52 I) Predictive pora X2 la distribución Conjunto de (f_x) es gaussiana, Condicionando se obtiene la predictiva nomo. Pf* (X*, t) = N(Nx, ox2) N= Kx (K+ On I) 1 + 52= K(Xx, Xx) - Kx (K+ 52 I) Kx Si desermes la predictiva subre tx : fx+ 17x anadimoi on a Kelacion Con Kernel ridge y modelo Boyesiano (mau) to media predictive de GPR: Un - Kx (K+ J, I) (Es retintion a la prediccion de kernel sidge (on 1=07

GPL y keinel ridge Comparten la misma media predictiva cuando se toma hiperparametros concidentes la difevencia principal es que GPR entrega una varianza Predictiva natural.

El GP pede vese Como el limite no parametrico del modelo lineal bayessiano cando el nuenero de Caracterifica Q fiende a infinito apropiodomente y el Piror indice el kernel k.