ΧΑΤΖΟΠΟΥΛΟΣ ΓΕΡΑΣΙΜΟΣ

ΔΥΙΚΟ ΠΡΟΒΛΗΜΑ ΚΑΙ ΕΦΑΡΜΟΓΗ ΣΤΗΝ R

1. Ας ξεκινήσω πρώτα από το δυικό πρόβλημα

$$\min_{x\in\mathbb{R}^d} \quad \frac{1}{2}\langle \boldsymbol{x}, \boldsymbol{Qx}\rangle + \langle \boldsymbol{q}, \boldsymbol{x}\rangle$$
$$\boldsymbol{Ax} \leq \boldsymbol{b},\ \boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{d}},\ \boldsymbol{b} \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{1}}$$

Και ας αποδείξω ότι το παραπάνω είναι ισοδύναμο με το ακόλουθο έτσι ώστε να ξεκινήσω να φτιάχνω τον αλγόριθμο μου στην R

$$\min_{\lambda\in\mathbb{R}^m} \frac{1}{2}\left\langle \boldsymbol{\lambda}, \boldsymbol{AQ^{-1}A^T\lambda}\right\rangle + \left\langle \boldsymbol{AQ^{-1}q} + \boldsymbol{b}, \boldsymbol{\lambda}\right\rangle + \frac{1}{2}\left\langle \boldsymbol{q}, \boldsymbol{Q^{-1}q}\right\rangle$$
$$\boldsymbol{\lambda} \geq \boldsymbol{0}$$

.Επίσης προσπαθείστε να αναλύσετε το ίδιο πρόβλημα χρησιμοποιώντας τον primal descent - dual ascent αλγόριθμο. Εφαρμόστε τους αλγόριθμους σε κάποιο πρόβλημα της επιλογής σας.

ΛΥΣΗ

Θέλω να κάνω Minimize $1/2\boldsymbol{x}^{\mathrm{T}}Q\boldsymbol{x} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ **(1)**

Ξέρω ότι η Lagrangian συνάρτηση είναι

$$L(x,\lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) = \frac{1}{2}x^{\mathrm{T}}\boldsymbol{Q}x + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \lambda^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{Ax})$$

το dual quadratic προβλημα είναι $*\boldsymbol{c}^{\mathrm{T}} = q$
$\mathrm{Max}_{\lambda\geq 0}\{\mathrm{Min}_x\, L(x,\lambda)\}$, i.e., $\mathrm{Max}_{\lambda\geq 0}\, L(x,\lambda)$

άρα
Maximize $L(\boldsymbol{x},\boldsymbol{\lambda}) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{Ax}) + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx}$ **(2)**

Οι διπλοί περιορισμοί υπονοούν ότι

$\boldsymbol{x}^{\mathrm{T}}[\boldsymbol{Qx} + \boldsymbol{c} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}] = \boldsymbol{x}^{\mathrm{T}}[\boldsymbol{0}]$
π.χ., $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{c} - \boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{c} - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{Ax} = 0$
The objective function can be rearranged as
$\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b} - 1/2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx} + (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{c} - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{Ax}) = \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b} - 1/2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx}$ from
όποτε το dual process μαζεύεται στο παρακατω
Maximize $\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b} - 1/2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx}$

με $Qx + c - A^{\mathrm{T}}\lambda = 0; \lambda \geq 0$.
για $Q = [\mathbf{0}]$, τότε η
(1) is $\mathrm{Min}\, \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$, με $\boldsymbol{Ax} \geq \boldsymbol{b}$
και η$(\mathbf{2})$ είναι $\mathrm{Max}\, \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b}$, με $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} = \boldsymbol{c}$

για $\boldsymbol{Q} \neq [\mathbf{0}]$
(1) έχει n μεταβλητές, $και$ m γραμμικούς περιορισμούς ανισότητας

(2) έχει $(m + n)$ μεταβλητές, n περιορισμούς ισότητας, and m μη αρνητικούς περιορισμούς
If $\boldsymbol{Q}$ is nonsingular (i.e., $\boldsymbol{Q}^{-1}$ exists) then from the dual constraints,

$$Qx + c - A^{\mathrm{T}}\lambda = \mathbf{0} \Rightarrow x + Q^{-1}c - Q^{-1}A^{\mathrm{T}}\lambda = \mathbf{0}$$

Άρα το $\boldsymbol{x} = \boldsymbol{Q}^{-1}[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}]$; όποτε πάμε να αντικαταστήσουμε το χ μας στο δυαδικό μας πρόβλημα

Ακόμα ξέρω αυτές τις 4 ιδιότητες για 2 πινάκες

i)$\boldsymbol{UV}$^T $= \boldsymbol{V}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}$, ii)$[\boldsymbol{U}^{\mathrm{T}}]^{\mathrm{T}} = \boldsymbol{U}$
και (iii) $U^{\mathrm{T}}\boldsymbol{V} = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{U}$ (αν είναι συμβατοί βέβαια). Also, (iv) $\boldsymbol{Q}, \boldsymbol{Q}^{-\mathrm{I}}$ είναι συμμετρικοί και ίδιοι με τους αντιπαραθετικούς (transpose).

Όποτε κάνουμε αντικατάσταση το χ που βρήκαμε πριν $\boldsymbol{x} = \boldsymbol{Q}^{-1}[\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}]$

$$\begin{aligned}\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b} - 1/2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Qx} &= \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{b} - 1/2\left[\boldsymbol{Q}^{-1}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}\right)\right]^{\mathrm{T}} \cdot \boldsymbol{Q} \cdot \left[\boldsymbol{Q}^{-1}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}\right)\right] \\ &= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}\right)^{\mathrm{T}} \cdot \boldsymbol{Q}^{-1} \cdot \boldsymbol{Q} \cdot \boldsymbol{Q}^{-1} \cdot \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}\right)\end{aligned}$$

Από i) και iv)

$$\begin{aligned}&= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2\left[\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}\right)^{\mathrm{T}} - \boldsymbol{c}^{\mathrm{T}}\right] \cdot \boldsymbol{Q}^{-1} \cdot \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}\right) \\ &= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2[\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}] \cdot \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}\right)\end{aligned}$$

Από i) και ii)

$$\begin{aligned}&= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2[\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{c} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1}\boldsymbol{c}] \\ &= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2[\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{c} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - (\boldsymbol{AQ}^{-1}\boldsymbol{c})^{\mathrm{T}}\boldsymbol{\lambda}] \\ &= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2[\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{c} - \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} - \boldsymbol{c}^{\mathrm{T}}(\boldsymbol{Q}^{-1})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}] \\ &= \boldsymbol{b}^{\mathrm{T}}\boldsymbol{\lambda} - 1/2[\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{c} - 2\boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}] \\ &= [\boldsymbol{b}^{\mathrm{T}} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{A}^{\mathrm{T}}]\boldsymbol{\lambda} - 1/2[\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{AQ}^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}] - 1/2[\boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}^{-1}\boldsymbol{c}]\end{aligned}$$

όμως ξερω ότι
$\min_{\lambda \in \mathbb{R}^m, \lambda \geq 0} L(x^*, \lambda) = -\max_{\lambda \in \mathbb{R}^m, \lambda \geq 0} L(x^*, \lambda)$

Άρα καταλήγουμε σε αυτό που θέλουμε να αποδείξουμε

$$[\boldsymbol{b}^T + \boldsymbol{c}^T\boldsymbol{Q}^{-1}\boldsymbol{A}^T]\boldsymbol{\lambda} + 1/2[\boldsymbol{\lambda}^T\boldsymbol{AQ}^{-1}\boldsymbol{A}^T\boldsymbol{\lambda}] + 1/2[\boldsymbol{c}^T\boldsymbol{Q}^{-1}\boldsymbol{c}]$$

Δηλαδή $1/2[\boldsymbol{\lambda}^T\boldsymbol{AQ}^{-1}\boldsymbol{A}^T\boldsymbol{\lambda}] + [\boldsymbol{b}^T + \boldsymbol{c}^T\boldsymbol{Q}^{-1}\boldsymbol{A}^T]\boldsymbol{\lambda} + \mathbf{1/2}[\boldsymbol{c}^T\boldsymbol{Q}^{-1}\boldsymbol{c}]$

Ακομα ας αναλυσω το ίδιο πρόβλημα με την βοηθεια του primal descent - dual ascent αλγόριθμο

Στην ουσία θέλω το min και το max της $L(x, \lambda$) με Ax<=b

Lagrange= $L(\boldsymbol{x},\boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})$

Όπου για το min θα πάρω το gradient δηλαδή την πρώτη παράγωγο ως προς χ της lagrange

Για το max θα πάρω την πρώτη παράγωγο ως προς το λ της lagrange

$$\begin{aligned} min = x^{(k+1)} \quad &= x^{(k)} - \alpha_k(\boldsymbol{c}^{\mathrm{T}} + \boldsymbol{\lambda}^T * A + Qx) \\ max = \lambda^{(k+1)} \quad &= \lambda^{(k)} + \alpha_k\left(Ax^{(k)} + b\right)_+ \end{aligned}$$

.

.

.

.

.

.

.

.

.

.

.

.

.

.

Με βαση τους παραπάνω αλγορίθμους που βρήκα ας τους εφαρμόσω σε ένα πρακτικό πρόβλημα που μπορώ να συναντήσω(quadratic lagrange minimize)

Ένα εργοστάσιο κατέληξε ότι το συνολικό τους κόστος για την παράγωγη προϊόντων όλου του μήνα βγαίνει από αυτήν την εξίσωση με χ1 το κόστος του πρώτου αγαθού και χ2 του δευτέρου αγαθού

Cost=0.4* X_1^2-5* X_1+ X_2^2 -6* X_2+50

Με τους ακολούθους περιορισμούς

$X_1 - X_2 \geq 2$ -> $X_1 - X_2 - 2 \geq 0$

$0.3 * X_1 + X_2 \geq 8$ -> $0.3 * X_1 + X_2 - 8 \geq 0$

$0 \leq X_1 \leq 10$

$0 \leq X_2 \leq 10$

Σκοπός μας είναι να βρούμε το μικρότερο δυνατό κόστος κάτω από αυτούς τους περιορισμούς για να πετύχουμε minimize του κόστους όπου θα αποφέρει περισσότερα έσοδα στην επιχείρηση

Άρα οι περιορισμοί μας σε πινακοειδής μορφή

$$\boldsymbol{Q} = \begin{bmatrix} \mathbf{0.4} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix},\ \boldsymbol{c}^T = \begin{bmatrix} \mathbf{-5} \\ \mathbf{-6} \end{bmatrix},\ \boldsymbol{A} = \begin{bmatrix} \mathbf{1} & \mathbf{-1} \\ \mathbf{-0.3} & \mathbf{-1} \end{bmatrix},\ \boldsymbol{b} = \begin{bmatrix} \mathbf{-2} \\ \mathbf{-8} \end{bmatrix}$$

.

.

.

.

CODE IN R

```
#οι περιορισμοί πρώτα σε πινακοειδής μορφή

Q <- matrix(c(0.4,0,0,1), nrow=2)

c <- matrix(c(-5,-6),ncol=1)

b <- matrix(c(-2,-8),ncol=1)

A <- matrix(c(1,-1,-0.3,-1), nrow=2)

#η lagrasian  συναρτηση μου

Lagrangian <- function(x,Q,q,A,b,l){

   L= t(c)%*%x+t(l)%*%A%*%x-t(l)%*%b+1/2*(t(x)%*%Q%*%x)#η (2)

   return(L)

}
```

```
#το positive projectuion
proj_pos_cone<-function(x)
{
  n=length(x)
  px=rep(0,n)
  for (i in 1:n){
    px[i]=max(x[i],0)
  }


  return(px)
}
proj_c<-function(x)
{ cone=proj_pos_cone(x)
return(cone)
}
gradient <- function(x,Q,c,A,b,l,a,max.iter){
  l1 <- l
  x1 <- x
  L_history <- c()
  for(i in 1:max.iter){
    x2 <- x1 - a*(Q%*%x1+c+t(A)%*%l1)#min παρ ως προς χ
    flambda<- l1 + a*(A%*%x1+b)#max παραγωγος ως λ
    l2 <- proj_c(flambda)
    x1 <- x2
    l1 <- l2
    L_history[i] <- Lagrangian (x1,Q,c,A,b,l1)
  }
```

```
   results <- list(x2,l2,L_history)

   return(results)

}

X0<- c(0,0) #init values

L0 <- c(0,0)#init values

a <- 0.1

max.iter <- 1000

results <- gradient(X0,Q,c,A,b,L0,a,max.iter)

results[1]
```

[,1]

[1,] 4.102317   #=x1 αγαθο

[2,] 7.007722   #=x2 αγαθο

#και οι 2 τιμές πληρούν πλήρως τις προϋποθέσεις μας

```
results[2]
```

[1] 3.359073 0.000000   #lambda

#άρα άμα κάνω αντικατάσταση τα χ1=4.102,χ2=7.007 που βρηκα θα βρω το ελαχιστο δυνατό κόστος που μπορεί να πετυχει το εργοστάσιο με βάση αυτά τα δυο αγαθα και με τους αντίστοιχους περιορισμους τους

Cost=0.4*x1^2-5*x1+x2^2-6*x2+50= 43.224

το ελάχιστο κόστος με αυτούς τους περιορισμούς σε αυτά τα αγαθά

ΕΥΧΑΡΙΣΤΩ