Duality

5.3

Lagrangian function:

$$L(x,\lambda) = c^T x + \lambda f(x)$$

Dual function:

$$\begin{aligned} g(\lambda) &= \inf_x \left(L(x,\lambda)\right) \\ &= \inf_x \left(c^T x + \lambda f(x)\right) \\ &= \begin{cases} -\lambda \sup_x (-\frac{c^T}{\lambda} x - f(x)), \lambda > 0 \\ -\infty, \lambda = 0 \end{cases} \\ &= \begin{cases} -\lambda f^*(-\frac{c}{\lambda}), \lambda > 0 \\ -\infty, \lambda = 0 \end{cases} \end{aligned}$$

Dual problem:

$$\begin{aligned} \max_x \quad & -\lambda f^*(-\frac{c}{\lambda}) \\ \text{s.t.} \quad & \lambda > 0 \end{aligned}$$

Since $h(c, -\lambda) = -\lambda f^*(-\frac{c}{\lambda})$ is the perspective function of convex function $f^*(c)$, the dual problem is convex.

5.4

(a) Let $c^T = \lambda_1 w^T A$, then:

$$\inf_x \left(c^T x\right) = \begin{cases} \lambda_1 w^T b, \lambda_1 \leq 0 \\ -\infty, \text{otherwise} \end{cases}$$

(b)

$$\begin{aligned} \max_{\lambda_1, w} \quad & \lambda_1 w^T b \\ \text{s.t.} \quad & w \succeq 0, \lambda_1 \leq 0 \\ & c = \lambda_1 A^T w \end{aligned}$$

(c) Lagrangian function:

$$L(x,\lambda_2) = c^T x + \lambda_2^T (Ax - b)$$

Dual function:

$$\begin{aligned} g(\lambda) &= \inf_x \left(c^T x + \lambda_2^T (Ax - b)\right) \\ &= \inf_x \left((c^T 1 + \lambda_2^T A)x - \lambda_2^T b\right) \\ &= \begin{cases} -\lambda_2^T b, (c^T 1 + \lambda_2^T A) = 0 \\ -\infty, \text{otherwise} \end{cases} \end{aligned}$$

Dual problem:

$$\begin{aligned}\max_{\lambda_2} \quad & -\lambda_2^T b \\ \text{s.t.} \quad & (c^T 1 + \lambda_2^T A) = 0 \\ & \lambda_2 \succeq 0\end{aligned}$$

Denote $\lambda^T = -\lambda_2^T = \lambda_1 w^T$, then the two problems above becomes the same:

$$\begin{aligned}\max_{\lambda} \quad & \lambda^T b \\ \text{s.t.} \quad & c = A^T\lambda, \lambda \preceq 0\end{aligned}$$

5.6

(a) Given the fact that $\|z\|_\infty \le \|z\|_2, \|z\|_2 \le \sqrt{m}\|z\|_\infty$:

$$\|Ax_{ls} - b\|_\infty \le \|Ax_{ls} - b\|_2 \le \|Ax_{ch} - b\|_2 \le \sqrt{m}\|Ax_{ch} - b\|_\infty$$

(b)

$$\begin{aligned} x_{ls} &= (A^TA)^{-1}A^Tb \\ \Rightarrow r_{ls} &= b - A(A^TA)^{-1}A^Tb \\ \Rightarrow A^Tr_{ls} &= A^Tb - A^TA(A^TA)A^Tb = 0 \\ \Rightarrow A^T\hat{\nu} &= A^T\tilde{\nu} = 0 \end{aligned}$$

and $\|\hat{\nu}\|_1 = \|\tilde{\nu}\|_1 = 1 \le 1$, hence they're both feasible.

$$b^T\tilde{\nu} = \frac{b^Tr_{ls}}{\|r_{ls}\|_1} = \frac{b^Tr_{ls} - ((A^Tr_{ls})^Tx_{ls})^T}{\|r_{ls}\|_1} = \frac{(b - Ax_{ls})^Tr_{ls}}{\|r_{ls}\|_1} = \frac{\|r_{ls}\|_2^2}{\|r_{ls}\|_1}$$

$$b^T\hat{\nu} = -b^T\tilde{\nu} = -\frac{\|r_{ls}\|_2^2}{\|r_{ls}\|_1} < b^T\tilde{\nu}$$

$b^T\tilde{\nu}$ is the better bound.

The bound in part (a) is $\frac{\|Ax_{ls}-b\|_\infty}{\sqrt{m}} = \frac{\|r_{ls}\|_\infty}{\sqrt{m}} \le \frac{\|r_{ls}\|_2}{\sqrt{m}} \le \frac{\|r_{ls}\|_2^2}{\|r_{ls}\|_1} = b^T\tilde{\nu}$.

5.7

(a) Lagrangian function:

$$L(x, \lambda) = \max_i (y_i) + \sum_i \lambda_i(a_i^Tx + b_i - y_i)$$

Dual function:

$$\begin{aligned} g(\lambda) &= \inf_{x,y}\left(\max_i (y_i) + \sum_i \lambda_i(a_i^Tx + b_i - y_i)\right) \\ &= \begin{cases} \inf_y \left(\max_i (y_i) + \lambda^Tb - \lambda^Ty\right), \sum_i \lambda_i a_i = 0 \\ -\infty, \text{otherwise} \end{cases} \\ &= \begin{cases} \lambda^Tb, \sum_i \lambda_i a_i = 0, \lambda \succeq 0, 1^T\lambda = 1 \\ -\infty, \text{otherwise} \end{cases} \end{aligned}$$

Dual problem:

$$\begin{aligned} \max_{\lambda} \quad & \lambda^T b \\ \text{s.t.} \quad & \sum_i \lambda_i a_i = 0 \\ & \lambda \succeq 0, 1^T\lambda = 1 \end{aligned}$$

(b)

$$\begin{aligned} \max_{x,y} \quad & y \\ \text{s.t.} \quad & a_i^T x + b_i - y \leq 0, i = 1, ..., m \end{aligned}$$

Dual function:

$$g(\lambda) = \inf_{x,y} \left(y + \sum_i \lambda_i (a_i^T x + b_i - y) \right)$$

Dual problem:

$$\begin{aligned} \max_{\lambda} \quad & \lambda^T b \\ \text{s.t.} \quad & \sum_i \lambda_i a_i = 0 \\ & \lambda \succeq 0, 1^T\lambda = 1 \end{aligned}$$

The LP dual is equal to the dual obtained in part (a).

(c) Let $b^T\nu^* - \sum_i \nu_i^* \log \nu_i^*$ be the optimal value of the dual GP, then

$$\begin{aligned} p_{gp}^* = b^T\nu^* - \sum_i \nu_i^* \log \nu_i^* \leq p_{pwl}^* - \sum_i \nu_i^* \log \nu_i^* \\ \Rightarrow p_{gp}^* - p_{pwl}^* \leq - \sum_i \nu_i^* \log \nu_i^* \leq \log m \end{aligned}$$

Therefore, $0 \leq p_{gp}^* - p_{pwl}^* \leq \log m$.

(d) Lagrangian function:

$$L(x, y, z) = \frac{1}{\gamma} \log \sum_i \exp(\gamma y_i) + z^T(Ax + b - y)$$

Dual function:

$$g(z) = \begin{cases} \inf_y \frac{1}{\gamma} \log \sum_i \exp(\gamma y_i) + z^T(b - y), A^T z = 0 \\ -\infty, \text{otherwise} \end{cases}$$

Let $\frac{d\frac{1}{\gamma} \log \sum_i \exp(\gamma y_i) + z^T(b-y)}{dy} = 0$, then we get $\frac{e^{\gamma y_i}}{\sum_i e^{\gamma y_i}} = z_i$ for all i.

That is solvable only when $1^T z = 1, z \succeq 0$.

At this moment, $y_i = \frac{\log(z_i \sum_j e^{\gamma y_j})}{\gamma}$.

$$\Rightarrow g(z) = \begin{cases} b^T z - \frac{1}{\gamma} \sum_i z_i \log z_i, A^T z = 0, 1^T z = 1, z \succeq 0 \\ -\infty, \text{otherwise} \end{cases}$$

Dual problem:

$$
\begin{aligned}
\max \quad & b^T z - \frac{1}{\gamma}\sum_i z_i \log z_i \\
\text{s.t.} \quad & A^T z = 0, 1^T z = 1
\end{aligned}
$$

Denote the optimal value of it as $p^*_{gp}(\gamma)$, then:

$$p^*_{gp}(\gamma) = b^T z^* - \frac{1}{\gamma}\sum_i z^*_i \log z^*_i \le b^T z^* + \frac{1}{\gamma}\log m = p^*_{pwl} + \frac{1}{\gamma}\log m$$

i.e. $p^*_{gp}(\gamma) - p^*_{pwl} \le \frac{1}{\gamma}\log m$

In conclusion, as we increase γ, $p^*_{gp}(\gamma)$ become closer to p^*_{pwl}.

5.9

(a)

$$\begin{bmatrix} \sum_k a_k a_k^T & a_i \\ a_i^T & 1 \end{bmatrix} = \begin{bmatrix} \sum_{k\ne i} a_k a_k^T & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} a_i \\ 1 \end{bmatrix} [a_i, 1] \succeq 0$$

Then the Schur complements of the top left block $1 - a_i^T(\sum_k a_k a_k^T)^{-1} a_i \ge 0$ i.e. $a_i^T X_{sim} a_i \le 1$.

In conclusion, X_{sim} is feasible.

(b) Let $\lambda = t1$, then the problem becomes

$$
\begin{aligned}
\max_t \quad & \log\det\left(\sum_i a_i a_i^T\right) + n\log t - mt + n \\
\text{s.t.} \quad & t \ge 0
\end{aligned}
$$

Denote $h(t) = \log\det\left(\sum_i a_i a_i^T\right) + n\log t - mt + n$. Let $h'(t) = 0$, we get $t = \frac{n}{m}$, the optimal value of t .

The dual objective at this λ is

$$g(\lambda) = \log\det\left(\sum_i a_i a_i^T\right) + n\log\frac{n}{m} = \log\det\left(\sum_i a_i a_i^T \cdot \left(\frac{n}{m}\right)^n\right)$$

This is equal to the primal objective value for $X = \left(\frac{m}{n}\right)^n X_{sim}$ but no more then the optimal value.

The volume of ellipsoid is proportional to $e^{-\frac{F}{2}}$, where F is the optimal value of the primal objective.

Therefore, the volume of the given ellipsoid is no more than $\left(\frac{m}{n}\right)^{\frac{n}{2}}$ more than the volume of the minimum volume ellipsoid.

5.10

(a) Let $X = \sum_i x_i v_i v_i^T$, then the problem becomes:

$$
\begin{aligned}
\max_x \quad & \log\det(X^{-1}) \\
\text{s.t.} \quad & X = \sum_i x_i v_i v_i^T \\
& x \succeq 0, 1^T x = 1
\end{aligned}
$$

Lagrangian function:

$$
\begin{aligned}
L(x,U,u,w) &= \log\det(X^{-1}) + \mathbf{tr}(UX) - \sum_i x_i v_i^T U v_i - u^T x + w(1^T x - 1) \\
&= \log\det(X^{-1}) + \mathbf{tr}(UX) - \sum_i x_i (v_i^T U v_i + u_i - w) - w
\end{aligned}
$$

Dual function:

$$
g(U,u,w) = \begin{cases} \inf_X \left(\log\det(X^{-1}) + \mathbf{tr}(UX) - w\right), v_i^T U v_i + u_i - w = 0 \text{ for } i \in [1,p] \\ -\infty, \text{otherwise} \end{cases}
$$

Denote $h(X) = \log\det(X^{-1}) + \mathbf{tr}(UX) - w$, then

$$
h'(X) = \frac{d\log\det(X^{-1})}{d\det(X^{-1})} \cdot \frac{d\det(X^{-1})}{dX^{-1}} \cdot \frac{dX^{-1}}{dX^T} + U\frac{d\,\mathbf{tr}X}{dX^T} = -X^{-1} + U
$$

Let $h'(X) = 0$, then we get $X = U^{-1}$.

$$
g(U,u,w) = \begin{cases} \log\det U + p - w, v_i^T U v_i + u_i - w = 0 \text{ for } i \in [1,p] \\ -\infty, \text{otherwise} \end{cases}
$$

Dual problem:

$$
\begin{aligned}
\max_{U,w} \quad & \log\det U + p - w \\
\text{s.t.} \quad & v_i^T U v_i \le w \text{ for } i \in [1,p]
\end{aligned}
$$

Simplified dual problem:

$$
\begin{aligned}
\max_{U'} \quad & \log\det U' + p\log p \\
\text{s.t.} \quad & v_i^T U' v_i \le 1 \text{ for } i \in [1,p]
\end{aligned}
$$

(b) Let $X = \sum_i x_i v_i v_i^T$, then the problem becomes:

$$
\begin{aligned}
\max_{x} \quad & \mathbf{tr}(X^{-1}) \\
\text{s.t.} \quad & X = \sum_i x_i v_i v_i^T \\
& x \succeq 0, 1^T x = 1
\end{aligned}
$$

Lagrangian function:

$$
\begin{aligned}
L(x,U,u,w) &= \mathbf{tr}(X^{-1}) + \mathbf{tr}(UX) - \sum_i x_i v_i^T U v_i - u^T x + w(1^T x - 1) \\
&= \mathbf{tr}(X^{-1}) + \mathbf{tr}(UX) - \sum_i x_i (v_i^T U v_i + u_i - w) - w
\end{aligned}
$$

Dual function:

$$
g(U,u,w) = \begin{cases} \inf_X \left(\mathbf{tr}(X^{-1}) + \mathbf{tr}(UX) - w\right), v_i^T U v_i + u_i - w = 0 \text{ for } i \in [1,p] \\ -\infty, \text{otherwise} \end{cases}
$$

Denote $h(X) = \mathbf{tr}(X^{-1}) + \mathbf{tr}(UX)$, then $h'(X) = -X^{-2} + U$.

Let $h'(X) = 0$, then we get $X = U^{-\frac{1}{2}}, (U \in S_{++}^p)$.

$$g(U,u,w)=\begin{cases}2\,\mathbf{tr}(U^{\frac{1}{2}})-w, v_i^TUv_i+u_i-w=0 \text{ for } i\in[1,p], U\succ 0\\ -\infty, \text{otherwise}\end{cases}$$

Dual problem:

$$\begin{array}{ll}\max\limits_{U,w} & 2\,\mathbf{tr}(U^{\frac{1}{2}})-w\\ \text{s.t.} & v_i^TUv_i\le w \text{ for } i\in[1,p]\\ & U\succ 0\end{array}$$

Simplified dual problem:

$$\begin{array}{ll}\max\limits_{U} & (\mathbf{tr}(U'^{\frac{1}{2}}))^2\\ \text{s.t.} & v_i^TU'v_i\le 1 \text{ for } i\in[1,p]\\ & U'\succ 0\end{array}$$

5.11

Let $y_i=A_ix+b_i$, then the problem becomes

$$\begin{array}{ll}\min\limits_{x,y} & \sum\limits_i\|y_i\|_2+\dfrac{\|x-x_0\|_2^2}{2}\\ \text{s.t.} & y_i=A_ix+b_i, i=1,...,N\end{array}$$

Lagrangian function:

$$L(x,\lambda_1,\lambda_2,...,\lambda_N)=\sum_i\|y_i\|_2+\frac{\|x-x_0\|_2^2}{2}+\sum_i\lambda_i^T(y_i-A_ix-b_i)$$

Dual function:

$$\begin{aligned}g(\lambda_1,\lambda_2,...,\lambda_N)&=\inf_{x,y}\left(\sum_i\|y_i\|_2+\frac{\|x-x_0\|_2^2}{2}+\sum_i\lambda_i^T(y_i-A_ix-b_i)\right)\\ &=\inf_{x,y}\left(\sum_i(\|y_i\|_2+\lambda_i^Ty_i)+(\frac{\|x-x_0\|_2^2}{2}-\sum_i\lambda_i^TA_ix)-\sum_i\lambda_i^Tb_i\right)\end{aligned}$$

$$\inf_y\left(\sum_i(\|y_i\|_2+\lambda_i^Ty_i)\right)=\inf_y\left(\sum_i\|y_i\|_2(1+\frac{\lambda_i^Ty_i}{\|y_i\|_2})\right)=\begin{cases}0, \|\lambda_i\|_2\le 1 \text{ for } i\in[1,N]\\ -\infty, \text{otherwise}\end{cases}$$

$$\inf_x\left(\frac{\|x-x_0\|_2^2}{2}-\sum_i\lambda_i^TA_ix\right)=\left(\frac{\|x-x_0\|_2^2}{2}-\sum_i\lambda_i^TA_ix\right)\Big|_{x=x_0+\sum_jA_j^T\lambda}=-\frac{\|\sum_iA_i^T\lambda_i\|_2^2}{2}-\sum_iA_ix_0^T\lambda_i$$

$$\Rightarrow g(\lambda_1,\lambda_2,...,\lambda_N)=\begin{cases}-\frac{\|\sum_iA_i^T\lambda_i\|_2^2}{2}-\sum_i(A_ix_0+b_i)^T\lambda_i, \|\lambda_i\|_2\le 1 \text{ for } i\in[1,N]\\ -\infty, \text{otherwise}\end{cases}$$

Dual problem:

$$\begin{array}{ll}\max\limits_{\lambda_1,\lambda_2,...,\lambda_N} & -\dfrac{\|\sum_iA_i^T\lambda_i\|_2^2}{2}-\sum\limits_i(A_ix_0+b_i)^T\lambda_i\\ \text{s.t.} & \|\lambda_i\|_2\le 1, i=1,...,N\end{array}$$

5.12

Let $y_i = b_i - a_i^T x$, then the problem becomes

$$\begin{aligned} \min_{x,y} \quad & -\sum_i \log(y_i) \\ \text{s.t.} \quad & y = b - Ax \end{aligned}$$

Lagrangian function:

$$L(x, y, \lambda) = -\sum_i \log(y_i) + \lambda^T(y - b + Ax)$$

Dual function:

$$\begin{aligned} g(\lambda) &= \inf_{x,y} \left(-\sum_i \log(y_i) + \lambda^T(y - b + Ax) \right) \\ &= \begin{cases} \sum_i \log \lambda_i + m - \lambda^T b, A^T\lambda = 0, \lambda \succ 0 \\ -\infty, \text{otherwise} \end{cases} \end{aligned}$$

Dual problem:

$$\begin{aligned} \max_{\lambda} \quad & \sum_i \log \lambda_i + m - \lambda^T b \\ \text{s.t.} \quad & A^T\lambda = 0, \lambda \succ 0 \end{aligned}$$

5.13

(a) Dual function:

$$\begin{aligned} g(\lambda, \nu) &= \inf_x \left(c^T x + \lambda^T(Ax - b) + \nu^T x - x^T \mathbf{diag}(\nu) x \right) \\ &= \inf_x \left(-x^T \mathbf{diag}(\nu) x + (c^T + \lambda^T A + \nu^T)x - \lambda^T b \right) \\ &= \begin{cases} \frac{1}{4} \sum_i \frac{(c_i + \lambda^T a_i + \nu_i)^2}{\nu_i} - \lambda^T b, \nu \preceq 0 \\ -\infty, \text{otherwise} \end{cases} \end{aligned}$$

Dual problem:

$$\begin{aligned} \max_{\lambda,\nu} \quad & \frac{1}{4} \sum_i \frac{(c_i + \lambda^T a_i + \nu_i)^2}{\nu_i} - \lambda^T b \\ \text{s.t.} \quad & \lambda \succeq 0, \nu \preceq 0 \end{aligned}$$

(b) Consider simplifying the dual problem in part (a):

Let $f(\nu_i) = \frac{(c_i+\lambda^T a_i+\nu_i)^2}{\nu_i}$, then $f(\nu_i) = \begin{cases} 4(c_i + \lambda^T a_i), c_i + \lambda^T a_i \le 0 \\ 0, c_i + \lambda^T a_i > 0 \end{cases} = \min\{0, 4(c_i + \lambda^T a_i)\}$.

Therefore, the dual problem can be simplified as:

$$\begin{aligned} \max_{\lambda} \quad & \sum_i \min\{0, (c_i + \lambda^T a_i)\} - \lambda^T b \\ \text{s.t.} \quad & \lambda \succeq 0 \end{aligned}$$

For the LP relaxation:

- Lagrangian function:

$$
\begin{aligned}
L(x,u,v,w) &= c^Tx + u^T(Ax-b) - v^Tx + w^T(x-1) \\
&= (c^T + u^TA - v^T + w^T)x - b^Tu - 1^Tw
\end{aligned}
$$

- Dual function:

$$
g(u,v,w) = \begin{cases} -b^Tu - 1^Tw, c^T + u^TA - v^T + w^T = 0 \\ -\infty, \text{otherwise} \end{cases}
$$

- Dual problem:

$$
\begin{aligned}
\max_{u,v,w} \quad & -b^Tu - 1^Tw \\
\text{s.t.} \quad & u \succeq 0, v \succeq 0, w \succeq 0 \\
& c^T + u^TA - v^T + w^T = 0
\end{aligned}
$$

- Another form of dual problem:

$$
\begin{aligned}
\max_{u,v,w} \quad & 1^T(c + A^Tu - v) - b^Tu \\
\text{s.t.} \quad & u \succeq 0, v \succeq 0 \\
& (c + A^Tu - v) \preceq 0
\end{aligned}
$$

As we can see, the dual problem of Lagrangian relaxation and the dual problem of LP relaxation are the same. So the lower bounds obtained via them are the same.

5.21

(a) The constraint set is $\{(x,y)|x=0, y>0\}$, which is convex.

The objective function $f(x,y)$ is convex since $\nabla^2 f = \begin{pmatrix} e^{-x} & 0 \\ 0 & 0 \end{pmatrix} \succeq 0$.

Hence this is a convex optimization problem. The optimal value is obviously 1.

(b) Dual function:

$$
g(\lambda) = \inf_{x,y>0}\left(e^{-x} + \lambda\frac{x^2}{y}\right) = \begin{cases} 0, \lambda \geq 0 \\ -\infty, \lambda < 0 \end{cases}
$$

Dual problem:

$$
\begin{aligned}
\max_{\lambda} \quad & 0 \\
\text{s.t.} \quad & \lambda \succeq 0
\end{aligned}
$$

Obviously, $\lambda^* \in \mathbb{R}, d^* = 0$, and the optimal duality gap is $p^* - d^* = 1$.

(c) Slater's condition does not hold for this problem.

(d) $p^*(u) = \begin{cases} 1, u = 0 \\ 0, u > 0 \end{cases}$.

When $u > 0$, the dual problem is:

$$
\begin{aligned}
\max_{\lambda} \quad & -\lambda u \\
\text{s.t.} \quad & \lambda \succeq 0
\end{aligned}
$$

Then $-\lambda^* u = 0$. So the global sensitivity inequality does not hold.

5.27

Lagrangian function:

$$
\begin{aligned}
L(x, \nu) &= \|Ax - b\|_2^2 + \nu^T (Gx - h) \\
&= x^T A^T A x + (G^T \nu - 2A^T b)^T x - \nu^T h
\end{aligned}
$$

KKT conditions:

1. Primal constraints: $Gx^* - h = 0$
2. Dual constraints: None
3. Complementary slackness: None
4. Gradient of Lagrangian with respect to x vanishes: $2A^T A x^* + G^T \nu - 2A^T b = 0$

According to the KKT conditions, we can solve the equations and get:

$$
\nu^* = 2(G(A^T A)^{-1} G^T)^{-1} (G(A^T A)^{-1} A^T b - h)
$$

$$
x^* = (A^T A)^{-1} \left(G^T (G(A^T A)^{-1} G^T)^{-1} (h - G(A^T A)^{-1} A^T b) + A^T b \right)
$$

5.30

Lagrangian function:

$$
L(X, \lambda) = \mathbf{tr}\ X - \log \det X + \lambda^T (Xs - y)
$$

KKT conditions:

1. Primal constraints: $X^* s = y, X^* \succ 0$
2. Dual constraints: None
3. Complementary slackness: None
4. Gradient of Lagrangian with respect to x vanishes: $I - (X^*)^{-1} + \frac{1}{2}(\lambda s^T + s \lambda^T) = 0$

According to the KKT conditions:

$$
\begin{aligned}
&(X^*)^{-1} = I + \frac{1}{2}(\lambda s^T + s\lambda^T) \\
\Rightarrow & s = (X^*)^{-1} y = y + \frac{1}{2}(\lambda + (\lambda^T y) s) \\
\Rightarrow & \lambda^T y + y^T y = 1 \\
\Rightarrow & \lambda = (1 + y^T y) s - 2y \\
\Rightarrow & X^{-1} = I + (1 + y^T y) s s^T - y s^T - s y^T \\
\Rightarrow & X^* = I + y y^T - \frac{1}{s^T s} s s^T
\end{aligned}
$$
