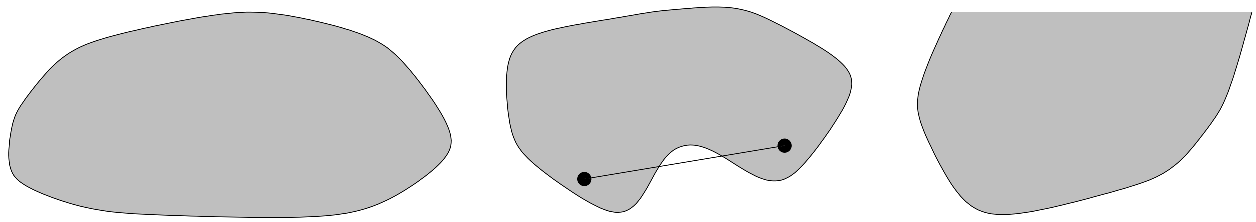
6 Linear Matrix Inequalities

6.1 Convex Sets

Definition: The set $\Omega \subseteq X$, where X is a real linear vector space, is convex if for any $x_1, x_2 \in \Omega$ and $\alpha \in [0, 1]$ the vector $x = \alpha x_1 + (1 - \alpha) x_2 \in \Omega$.



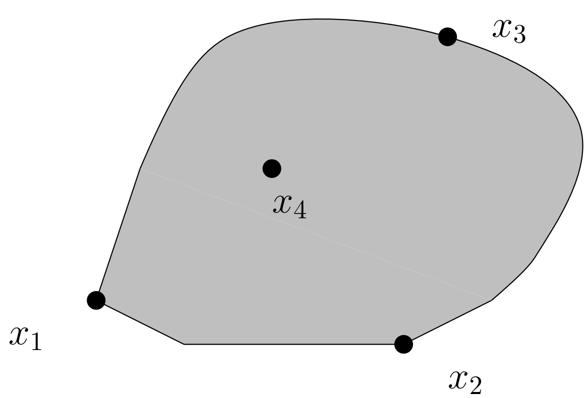
convex   not convex   convex

Convex combination: Given the convex set Ω and the vectors $x_i \in \Omega$, $i = 1, \ldots, N$, then

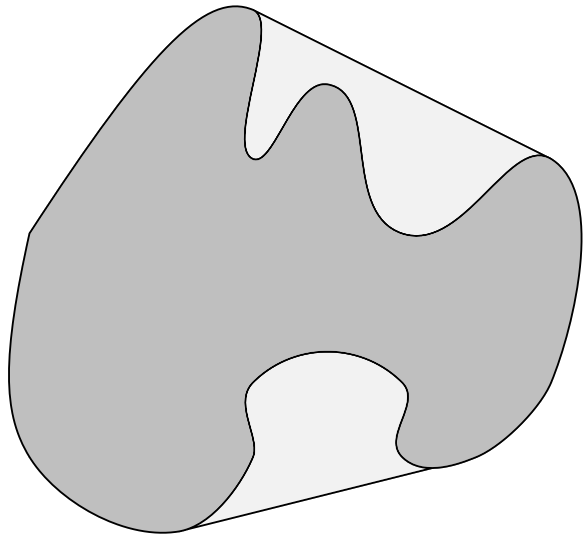
$$x = \sum_{i=1}^{N} \xi_i x_i \in \Omega, \qquad \text{for all } \xi_i \geq 0, \qquad \sum_i \xi_i = 1.$$

Extreme Points (x_1, x_2, x_3): A vector $x \in \Omega$, where Ω is a convex set, is an *extreme point* if it cannot be generated by a convex combination of any other vector in Ω.

Interior Points (x_4): A vector $x \in \Omega$, where Ω is a convex set, is an *interior point* if it is not an extreme point.



Convex Hull: Given an arbitrary set Γ, its *convex hull* is the smallest convex set containing Γ. (Notation: $\mathrm{co}(\Gamma)$)



Proposition: Let Ω and Γ be convex sets.

a) $\alpha\Omega := \{x : x = \alpha x, x \in \Omega\}$ is a convex set.

b) $\Omega + \Gamma := \{z : z = x + y, x \in \Omega, y \in \Gamma\}$ is a convex set.

c) $\Omega \cap \Gamma$ is a convex set.

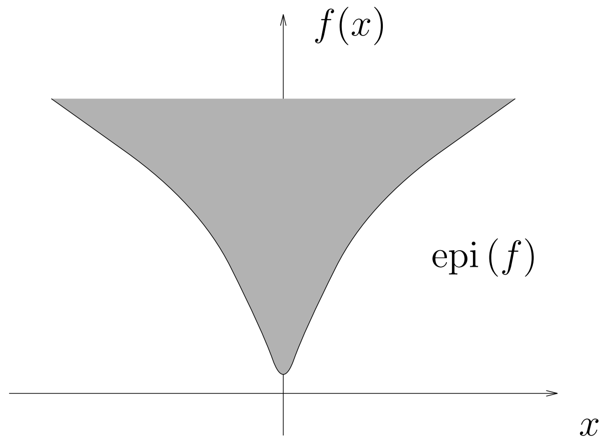
6.2 Convex Functions

Definition: The real valued function $f : \Omega \to \mathbb{R}$, where $\Omega \subseteq X$ is a convex set, is said to be a *convex function* if
$f(\alpha x_1 + (1-\alpha)x_2)) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$, $\forall x_1, x_2 \in \Omega$, $\alpha \in [0,1]$.

Epigraph: Given a function $f : \Omega \to \mathbb{R}$ the set

$$\text{epi}\,(f) := \{(x,\gamma) : x \in \Omega, \gamma \in \mathbb{R}, f(x) \leq \gamma\}$$

is the *epigraph of* f.



Convex function: The function $f : \Omega \to \mathbb{R}$ is convex iff $\text{epi}\,(f)$ is a convex set.

Proposition: All convex functions are continuous.
Proposition: All norms are convex functions.

Concave function: The function $f : \Omega \to \mathbb{R}$ is concave iff $(-f)$ is convex.

Convexity (twice differentiable real function): Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. The following are equivalent

a) f is convex.

b) $f(y) \geq f(x) + \nabla f(x)(y - x) \quad \forall x, y \in \mathbb{R}^n$.

c) $h^T \nabla^2 f(x) h \geq 0 \quad \forall x, h \in \mathbb{R}^n$.

6.3 Cones

Definition: The set $\Gamma \subset X$, where X is a real vector space is a *cone with vertex at the origin* if $\alpha x \in \Gamma$ for all $x \in \Gamma$, $\alpha \geq 0$.

Convex Cone: $\Gamma \subset X$ is a *convex cone* if it is a convex cone. :)

Examples (Convex cones)

1. $\Gamma := \{x : x \in \mathbb{R}^n, \quad x_i \geq 0, \quad i = 1, \ldots, n\}$.

$$x, y \in \Gamma \;\Rightarrow\; \begin{cases} \forall \alpha \geq 0, & \alpha x_i \geq 0 \;\Rightarrow\; \alpha x \in \Gamma. \\ \forall \alpha \in [0,1], & \alpha x_i + (1-\alpha) y_i \geq 0 \;\Rightarrow\; \alpha x + (1-\alpha) y \in \Gamma. \end{cases}$$

2. $\Gamma := \{X : X \in \mathbb{S}^n, \quad x^T X x \geq 0, \, \forall x \in \mathbb{R}^n\}$.

$$X, Y \in \Gamma \;\Rightarrow\; \begin{cases} \forall \alpha \geq 0, & \alpha x^T X x \geq 0, \, \forall x \in \mathbb{R}^n \;\Rightarrow\; \alpha X \in \Gamma. \\ \forall \alpha \in [0,1], & \alpha x^T X x + (1-\alpha) x^T Y x \geq 0 \;\Rightarrow\; \alpha X + (1-\alpha) Y \in \Gamma. \end{cases}$$

Example (Nonconvex cone)

3. $\Gamma := \{X : X \in \mathbb{R}^{n \times n}, \quad \mathbb{R}\{\lambda_i(X)\} \leq 0, \, \forall i\}$.
   Remembering that $\lambda(\alpha X) = \alpha \lambda(X)$

$$X \in \Gamma \;\Rightarrow\; \forall i, \alpha \in \mathbb{R}, \alpha \geq 0, \quad \alpha \mathbb{R}\{\lambda_i(X)\} = \mathbb{R}\{\lambda_i(\alpha X)\} \leq 0 \;\Rightarrow\; \alpha X \in \Gamma.$$

For

$$X_1 = \begin{bmatrix} -1 & 4 \\ 0 & -1 \end{bmatrix}, \quad X_2 = X_1^T = \begin{bmatrix} -1 & 0 \\ 4 & -1 \end{bmatrix},$$

notice that $\lambda_i(X_1) = \lambda_i(X_2) = -1$, $i = 1, 2 \quad \Rightarrow \quad X_1, X_2 \in \Gamma$. However, for

$$X = \frac{1}{2} X_1 + \frac{1}{2} X_2 = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix},$$

$\lambda_1(X) = -3, \lambda_2(X) = 1 \quad \Rightarrow \quad X \notin \Gamma$.

6.4 Convex Functionals

Comparing vectors: Let Γ be a *convex cone* defined in X. Given $x, y \in X$ we say that $x \succeq y$ if $(x - y) \in \Gamma$ (Particular case: $x \succeq 0 \Leftrightarrow x \in \Gamma$). The convex cone Γ is usually referred to as a *positive cone*.

Definition: Let X and Z be vector spaces with $F = \mathbb{R}$. Assume that Z is equipped with a positive cone Γ. The mapping $f : \Omega \to Z$, where $\Omega \subseteq X$ is a convex set, is said to be *convex* if
$f(\alpha x_1 + (1-\alpha)x_2)) \preceq \alpha f(x_1) + (1-\alpha) f(x_2)$, $\forall x_1, x_2 \in X$, $\alpha \in [0, 1]$.

Proposition: If the functional $f : \Omega \to Z$ is convex then the set

$$\{x : x \in \Omega, f(x) \preceq y\}$$

is a convex set for all $y \in Z$.

Examples:

1. $f : \mathbb{S}^n \to \mathbb{S}^n$, $f(X) := A^T X + XA + Q$.
2. $f : \mathbb{S}^n \to \mathbb{S}^n$, $f(X) := XRX - A^T X - XA - Q$.
3. $f : \Omega \times \mathbb{R}^{m \times n} \to \mathbb{S}^m$, $\Omega := \{X : X \in \mathbb{S}^n, X \succ 0\}$, $f(X) := YX^{-1}Y^T$.
4. $f : \Omega \times \mathbb{R}^{n \times n} \to \mathbb{S}^m$, $\Omega := \{(X, Z) : X, Z \in \mathbb{S}^n, X \succ 0, Z \succ YX^{-1}Y^T\}$,
$$f(X) := Z + (Y - X)\left(Z - YX^{-1}Y^T\right)^{-1}(Y - X)^T.$$

6.5 Positivity for Symmetric Matrices

$\mathbb{S}^n$: symmetric matrices of dimensions $n \times n$

Γ: Positive Cone on $\mathbb{S}^n$

$$\Gamma := \{X : X \in \mathbb{S}^n, \quad x^T X x \geq 0, \forall x \in \mathbb{R}^n\},$$
$$\operatorname{int}\Gamma = \{X : X \in \mathbb{S}^n, \quad x^T X x > 0, \forall x \in \mathbb{R}^n, x \neq 0\}.$$

Positiveness on $\mathbb{S}^n$:
Given $X \in \mathbb{S}^n$ the following statements are equivalent:

a) X is *positive semidefinite*.

b) $X \in \Gamma \quad (X \succeq 0)$.

c) $\lambda_{\min}(X) \geq 0 \quad (\lambda_{\max}(-X) \leq 0)$.

a) X is *positive definite*.

b) $X \in \operatorname{int}\Gamma \quad (X \succ 0)$.

c) $\lambda_{\min}(X) > 0 \quad (\lambda_{\max}(-X) < 0)$.

Negativeness on $\mathbb{S}^n$:
A matrix $X \in \mathbb{S}^n$ is *negative* (*semi*)*definite* if $(-X \succeq 0)$ $-X \succ 0$;
Notation: $X \prec 0$ $(X \preceq 0)$.

6.6 Linear Matrix Inequalities

Affine functional:
The functional $f : X \to \mathbb{S}^n$ is *affine* if

$$f(\alpha x + (1-\alpha)y) = \alpha f(x) + (1-\alpha) f(y), \quad \forall x, y \in X, \alpha \in \mathbb{R},$$

Proposition: The set $\{x \,:\, f(x) \preceq 0\}$ (or $\{x \,:\, f(x) \prec 0\}$), where $f : X \to \mathbb{S}^n$ is an affine functional, is a convex set.

Proposition: The set $\{x \,:\, f(x) \succeq 0\}$ (or $\{x \,:\, f(x) \succ 0\}$), where $f : X \to \mathbb{S}^n$ is an affine functional, is a convex set.

Proof: An affine functional is simultaneously convex and concave.

Definition: Let $f : X \to \mathbb{S}^n$ be an affine functional. The functional inequalities

$$f(x) \preceq 0 \quad \text{or} \quad f(x) \succeq 0$$

are known as *Linear Matrix Inequalities* (LMI). The functional inequalities

$$f(x) \prec 0 \quad \text{or} \quad f(x) \succ 0$$

are *strict* LMIs.

Examples:

1. $f : \mathbb{S}^n \to \mathbb{S}^n, \quad f(X) := A^T X + XA + Q \preceq 0.$
2. $f : \mathbb{S}^n \to \mathbb{S}^n, \quad f(X) := A^T X + XA \prec 0.$
3. $f : \mathbb{S}^n \to \mathbb{S}^n, \quad f(X) := A^T XA - X + Q \preceq 0.$
4. $f : \mathbb{S}^n \to \mathbb{S}^{2n}, \quad f(X) := \begin{bmatrix} X & A^T X \\ XA & X \end{bmatrix} \succ 0.$

6.7 Semidefinite Programming (SDP)

Proposition: Any affine functional $F : X \to \mathbb{S}^n$ where X is a real *finite dimensional space* can be put in the canonical form

$$F(x) \succeq 0, \quad F(x) := F_0 + \sum_{i=1}^{m} F_i x_i, \quad F_i \in \mathbb{S}^n, x \in \mathbb{R}^m.$$

Examples:

1. $G(X) := A^T X + XA + Q \preceq 0, \quad X \in \mathbb{S}^2.$

$$X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} x_3 = \sum_{i=1}^{3} X_i x_i, \quad x \in \mathbb{R}^3.$$

$$\implies \quad -G(X) = \underbrace{-Q}_{F_0} + \sum_{i=1}^{3} \underbrace{-(A^T X_i + X_i A)}_{F_i} x_i = F(x)$$

2. $G(\mu, X) := \mu - \operatorname{trace}\left(CXC^T\right) \succeq 0, \quad \mu \in \mathbb{R}, X \in \mathbb{S}^2.$

$$G(\mu, X) = \mu - \operatorname{trace}\left(CXC^T\right) = \underbrace{1}_{F_i} \mu + \sum_{i=1}^{3} \underbrace{- \operatorname{trace}\left(CX_iC^T\right)}_{F_i} x_i = F(\mu, x)$$

Canonical Semidefinite Program:

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \quad & c^T x \\ \text{s.t.} \quad & F(x) \succeq 0 \end{aligned}$$

Remarks:

a) Several LMI $F_j(x) \succeq 0$, $j = 1, \ldots, N$ can be handled by stacking $F_j(x)$ as a block diagonal constraint

$$F(x) = \text{blockdiag}\left(F_1(x), \ldots, F_j(x), \ldots, F_N\right) \succeq 0.$$

b) Some LMI may have equality constraints (LME). Examples

1. $f : \mathbb{R}^{n \times n} \to \mathbb{S}^n, \quad f(X) := A^T X + XA + Q \preceq 0, \quad X = X^T.$
2. $f : \mathbb{R}^{n \times n} \to \mathbb{S}^n, \quad f(X) := A^T X + XA \prec 0, \quad CX = B^T.$

c) Use a parser to generate the F_is, e.g. YALMIP.

d) WARNING: Do not use Matlab's LMIToolbox :(

6.8 SDP with equality constraints

Canonical SDP with equality constraints:

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \quad & c^T x \\ \text{s.t.} \quad & F(x) \succeq 0 \\ & Ax = b \end{aligned}$$

Conceptual elimination of equality constraints: Solve $Ax = b$

$$\tilde{x}(y) = A^\dagger b + A^\perp y, \qquad y \in \mathbb{R}^p \qquad (p = \dim A^\perp)$$

then

$$\begin{aligned} c^T \tilde{x}(y) &= \tilde{c}_0 + \tilde{c}^T y, \qquad \tilde{c}_0 := c^T A^\dagger b, \qquad \tilde{c} := A^{\perp T} c, \\ \tilde{F}(y) &:= F(\tilde{x}(y)), \end{aligned}$$

Notice that

$$\min_{x \in \mathbb{R}^n, Ax = b} c^T x = \min_{y \in \mathbb{R}^p} c^T \bar{x}(y) = \min_{y \in \mathbb{R}^p} \left(\tilde{c}_0 + \tilde{c}^T y \right) = \tilde{c}_0 + \min_{y \in \mathbb{R}^p} \tilde{c}^T y$$

This provides the standard SDP

$$\begin{aligned} \min_{y \in \mathbb{R}^p} \quad & \tilde{c}^T y \\ \text{s.t.} \quad & \tilde{F}(y) \succeq 0. \end{aligned}$$

A good solver/parser will do that for you!

6.9 Some useful tools

6.9.1 Congruence Transformation

Definition: Two matrices $X, Y \in \mathbb{S}^n$ are said to be *congruent* if there exists a nonsingular matrix $T \in \mathbb{R}^{n\times n}$ such that $Y = T^T X T$.

Proposition: If X and Y are congruent then $Y \succ 0$ if, and only if, $X \succ 0$.

Proof: If $X \succ 0$ then $x^T X x > 0$, $\forall x \in \mathbb{R}^n, x \neq 0$. Since X and Y are congruent there exists T nonsingular such that $Y = T^T X T$. Hence, using the fact that T is nonsingular, for all $x \neq 0$, the vector $y := T^{-1}x \neq 0$ and

$$X \succ 0 \quad \Leftrightarrow \quad x^T X x = y^T T^T X T y = y^T Y y > 0 \quad \Leftrightarrow \quad Y \succ 0$$

Generalization: For any $T \in \mathbb{R}^{m\times n}$, the matrix $Y = T^T X T \succeq 0$ if $X \succeq 0$. The converse is false when T does not have full row rank; example: $t = 0, x = -1$, $y = t^2 x \geq 0$.

6.9.2 Schur Complement

Lemma: For all $X \in \mathbb{S}^n, Y \in \mathbb{R}^{m\times n}, Z \in \mathbb{S}^m$, the following statements are equivalent:

a) $Z \succ 0, \quad X - Y^T Z^{-1} Y \succ 0.$	a) $Z \succ 0, \quad X - Y^T Z^{-1} Y \succeq 0.$
b) $\begin{bmatrix} X & Y^T \\ Y & Z \end{bmatrix} \succ 0.$	b) $Z \succ 0, \quad \begin{bmatrix} X & Y^T \\ Y & Z \end{bmatrix} \succeq 0.$

Proof: Assume $Z \succ 0$. The *nonsingular matrix*

$$T = \begin{bmatrix} I & 0 \\ -Z^{-1}Y & I \end{bmatrix}$$

establishes the congruence transformation

$$T^T \begin{bmatrix} X & Y^T \\ Y & Z \end{bmatrix} T = \begin{bmatrix} X - Y^T Z^{-1} Y & 0 \\ 0 & Z \end{bmatrix} \succ 0 (\succeq 0).$$

6.10 Scope of Semidefinite Programming

Problem: Linear Programming

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq b \end{aligned}$$

Equivalent SDP:

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \quad & c^T x \\ \text{s.t.} \quad & F(x) \succeq 0, \end{aligned}$$

where $F(x) := \operatorname{blockdiag}(A_i x - b_i)$, A_i is the ith row of A.

Problem: Quadratic Programming (convex)

$$\begin{aligned} \min_{x \in \mathbb{R}^m} \quad & x^T Q^T Q x + c^T x \\ \text{s.t.} \quad & Ax \succeq b \end{aligned}$$

Using Schur complement

$$\gamma \geq x^T Q^T Q x + c^T x \quad \Leftrightarrow \quad F_Q(\gamma, x) := \begin{bmatrix} \gamma - c^T x & x^T Q^T \\ Qx & I \end{bmatrix} \succeq 0$$

Equivalent SDP:

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^m} \quad & \gamma \\ \text{s.t.} \quad & F(x) \succeq 0, \end{aligned}$$

where $F(\gamma, x) = \operatorname{blockdiag}(F_Q(\gamma, x), A_i x - b_i)$, A_i is the ith row of A.

Problem: Minimum maximum eigenvalue problem

$$\begin{aligned} \min_{x\in\mathbb{R}^m} \max_i \quad & \lambda_i(F(x)) \\ \text{s.t.} \quad & F(x) \in \mathbb{S}^n \end{aligned}$$

For any $X \in \mathbb{S}^n$

$$\min_i \lambda_i(X) I \preceq X \preceq \max_i \lambda_i(X) I$$

hence any γ such that

$$\gamma I \succeq F(x)$$

is also such that $\gamma \geq \max_i \lambda_i(F(x))$.

Equivalent SDP:

$$\begin{aligned} \min_{\gamma\in\mathbb{R}, x\in\mathbb{R}^m} \quad & \gamma \\ \text{s.t.} \quad & \gamma I \succeq F(x). \end{aligned}$$

Problem: Minimum Matrix Norm

$$\min_{X\in\mathbb{R}^{m\times n}} \quad \|A - BXC\|^2, \quad \text{where } \|M\| := \lambda_{\max}^{1/2}[M^T M]$$

Notice that

$$\gamma \geq \lambda_{\max}\left[(A - BXC)^T (A - BXC)\right] \quad \Leftrightarrow \quad \gamma I \succeq (A - BXC)^T (A - BXC)$$

Using Schur complement

$$\gamma I \succeq (A - BXC)^T (A - BXC) \Leftrightarrow F(\gamma, X) := \begin{bmatrix} \gamma I & (A - BXC)^T \\ (A - BXC) & I \end{bmatrix} \succeq 0.$$

Equivalent SDP

$$\begin{aligned} \min_{\gamma\in\mathbb{R}, X\in\mathbb{R}^{m\times n}} \quad & \gamma \\ \text{s.t.} \quad & F(\gamma, X) \succeq 0, \end{aligned}$$

6.11 Feasibility Semidefinite Program

Problem: Find $x \in \mathbb{R}^m$ such that $F(x) \succeq 0$.

Solution: Introduce the scalar $\gamma \in \mathbb{R}$ and solve the *Optimization* SDP

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^m} \quad & \gamma \\ \text{s.t} \quad & F(x) + \gamma I \succeq 0 \end{aligned} \tag{13}$$

Given x_0 not feasible then compute $\max_i \lambda_i(F(x_0))I$. Then

$$\mu_0 := \min_i \lambda_i(F(x_0)) - 1 \qquad \Longrightarrow \qquad \mu_0 I \prec \min_i \lambda_i(F(x_0)) \preceq F(x_0)$$

Therefore $\gamma_0 = -\mu_0$, x_0 are such that

$$F(x_0) + \gamma_0 I \succ 0$$

i.e., (γ_0, x_0) are strictly feasible solutions to the SDP. Most algorithms for SDP (interior-point methods) require a strictly feasible initial solution.

Let $(\tilde{\gamma}, \tilde{x})$ be the optimal solution to SDP (13). Then

a) If $\tilde{\gamma} > 0$ then there is no feasible solution to the original problem.

b) If $\tilde{\gamma} = 0$ then $\tilde{x}$ is a feasible solution to the original problem ($F(x) \succeq 0$).

c) If $\tilde{\gamma} < 0$ then $\tilde{x}$ is a *strictly feasible* solution to the original problem ($F(x) \succ 0$).